Dear Editor, Reviewer 1, and Reviewer 2:

Thank you for taking the time to provide all of your meaningful and insightful comments and suggestions. We have taken them all into serious consideration and have strived to work hard to address them all. In this response, your original comments are completely unedited and are given in yellow highlight, our responses are given in blue highlight and updates to the paper [redacted]. We respond to Reviewer 1 first in full, and then respond to Reviewer 2 in full. If answers are given above to a previous question, we may refer the reader to "see above" or something similar. Thank you again for your time and deep insights!

Response to Author 1:

Summary-
The paper compares a simple plume model and a multiple linear regression (MLR) model approach to observed plume heights from MISR. The plume model and MLR models use overlapping data sets to predict plume height. The authors find that the plume model generally under performs the MLR models. The use of overlapping data to train the MLR models and to get predictions from the plume model is interesting. However, the use of a single plume model from 1965 is poorly motivated. The authors need to discuss in detail the current state of the field in plume modelling (which I feel must have progressed somewhat in the past 50 years) and compare several plume models to the MLR models. At some level the MLR model will always get a better agreement with data because mathematically it is going to always minimize unexplained variance, in contrast to the plume model, which is based on some physical understanding.

This is an essential and important part of the paper that we have made modifications to make clearer. Thank you for pointing out this essential communication issue!

One of the critical assumptions of all plume rise models is that the vertical rise is controlled by the buoyancy and vertical motion forces. The input from the fire is the heat co-emitted with the aerosols and gasses. Any initial vertical momentum applied at the fire start point and the atmospheric temperature distribution are a function of the atmospheric state. As the heated air from the fire rises through the atmospheric column, it interacts with the background conditions and eventually an equilibrium state is reached.

However, there are a few factors which have been found to be important, but are missing in this approximation. There are now more than a few papers (Guo et al., 2019; Tao et al., 2012; and Mims et al., 2010) that show the aerosols co-emitted with the heat absorb and scatter a significant amount of incoming solar radiation in the daytime and outgoing IR radiation in the nighttime, changing the energy structure of the column above and below the point at which the aerosols are located in the vertical, as well and the buoyancy of the air parcel containing the aerosols. Secondly, in the case where there is a large-scale aerosol cloud due to extensive burning over a significant land surface area, this widely distributed cloud of aerosols in the atmosphere further changes the absorption and scattering of the atmosphere at the meso-scale (Wang et al., 2009; Ekman et al., 2011; Cohen et al., 2011), in turn further changing the atmosphere's general energy balance. A third issue is the radiative-convective equilibrium occurring within the column over which the air parcels rise also depend on the loadings of clouds and aerosols above and below the parcel of interest. Therefore, any physically-based plume-rise model, as currently found in the literature and used by the modeling communities, regardless of whether it was fitted 50 years ago or has been slightly improved in terms of its coefficients under different conditions, still cannot capture the required set of physics to be fully realistic. Hence, we do not feel that the issue is how long ago the currently used theory was developed is overly relevant. In fact, we attempt to form a regression model (as you term “MLR model” or simply “MLR” from this point forward) specifically to cater to this assumption (more on this later).

The following paragraph has been added into the paper in Section 3

Third, the range of the seven regression models is an attempt to intelligently account for the fact that the column loadings of the CO and NO_2 offer physical meaning and insight, as compared to merely being an attempt to minimize any unexplained variance. We argue that the column values of both CO and NO_2 are both directly and indirectly related to the magnitude and the height of the vertical aerosol column. Due to the fact that the emissions of NO_2 is a strong function of the fire temperature, and its short atmospheric lifetime, the NO_2 is strongly related to the temperature of the fire, or the FRP, which is one of the essential driving forces of the buoyancy. This issue is strongly coupled with the fact that FRP is also one of the most error-prone of the measurements commonly used to drive the plume-rise models, with the FRP commonly underestimated in the tropics due to clouds and aerosols, as given in Kaiser et al. (2012), Cohen et al. (2018), and Lin et al. (2020a). Additionally, the amount of CO produced is a function of the total amount of biomass burned as well as the wetness of the surface itself where the burning occurred, and hence the CO column loading is also physically related to the properties of the fires. In fact, using a measure of the CO column can help us to overcome the physical constraints that current measurements have in terms of addressing the issues of how much peat or understory has burned, or if such fires which are occurring without direct line of sight from above can even be detected by the current fire detection processes at all (Leung et al., 2007; Ichoku et al., 2008). The combination of high NO_2 (which is more produced at higher temperature) and low CO (which is more produced at higher temperature) means that the ratio of NO_2 to CO also provides further physical insight into the non-linearities associated with the fire temperature, wetness, and possibility of other heat sources/sinks at the fire/atmosphere interface such as smoldering, conversion to latent heat, etc.

This overfitting problem could be solved by training the MLR model in one region and applying it to other regions. The authors also train 7 MLR models based on a combination of different predictors. The way that this feeds into the comparisons between the ‘regression model’ and plume model is poorly described. The authors need to either use all the predictors, or come up with some objective methodology to throw out some (eg machine learning).

This is an interesting point, and I believe worthwhile for follow-up work. It is not well known if such a single model would allow for a single idealized modeling format to be achieved throughout the entire real world for three different reasons. First, the biomass type and loading are different across different regions of the world. Secondly, the climatology of the soil moisture, boundary layer, and the free atmospheric vertical profile are also not consistent across different parts of the world. Finally, these different regions are sometimes impacted by human emissions and sources of co-emitted heat, aerosols and gasses, and sometimes not. For this reason, in this work we are focusing first and foremost on the idea that applying a physically based MLR model can give us insights, and to figuring out where such approach may add value for the community as a whole.

The results show clearly that the versions of the regression model that best model the height all have the NO_2 term in them. Furthermore, over all regions except for one, the best fitting regression models also have a term representing CO. Therefore, the number of regression models computed, in retrospect, could have been reduced, with models 5, 6, and 7 excluded. This end result shows clearly that the simpler plume rise regression model representation is never superior in any case. Further work could look into how more advanced modeling perspectives may or may not improve upon the framework introduced here. We believe that there is already considerable value and uniqueness offered by this approach.

The following paragraph has been added into the paper in section 3.3

The regression model solely containing NO_2 is an approximation of the concept that the heat of the biomass burning should have an important role to play in terms of the plume height. Furthermore, using NO_2 in this way helps to get around the inherent underestimation of FRP. The regression model solely containing the CO is a proxy for the concept that the mass of biomass burned should make an important contribution towards the plume height. Inclusion of the CO term is also a way to get around the underapproximation of the total burned area, or of any significant contribution from underground burning.

The following sentences have been added into the paper in section 3.3

The regression model with the non-linear combination of the two is a proxy for the argument that it is the ratio of the heat to the total biomass burned that is an essential physical consideration to take into effect. Furthermore, this final case provides some weight to the concept that a small change in the vertical column concentration may have a stronger than linear effect, as is evidenced by (Ichoku et al., 2008; Zhu et al., 2018), such as in terms of absorbing aerosols (which are themselves produced more so under hot or oxygen starved conditions) in the vertical column altering the ultimate vertical distribution.

The following paragraph has been added into the paper in Section 2.7

The 7 different regression models were chosen so as to cover the entire combination of different ways to fairly and uniformly incorporate the CO and NO_2 measurements as well as their underlying physical meanings. The 7th regression model is the approximation of the Plume Rise Model. The 4th and 5th regression models are the approximations of the single-species linear impact of NO_2 and CO respectively. The 6th regression model approximates the single-species non-linear impact of NO_2 and CO in tandem. Finally, the 1st through 3rd regression models are the approximations of the combination of CO and NO_2 in tandem with both linear (model 1), or with one linear and one non-linear combination (models 2 and 3). This approach is consistent with and follows from some of the earlier works which tries to use advanced learning to understand some higher order, simple non-linear forcings, still based on some physical consideration, i.e. Cohen and Prinn, 2011.

The following sentence has been added into the paper in Section 4

As we have demonstrated, the impact of NO_2 (as a proxy for the burning temperature) is always essential, and the impact of CO (as a proxy for the total biomass burned) is usually essential as well. We further have shown that the simplest regression model, the approximation of the Plume Rise Model, never yields the best fit to the data.

The paper seems rushed and has many grammatical errors. The number of figures must be increased to make it clearer what the analysis shows.

We have included 3 new figures (Figures 4, 5, and 6) in the paper and 2 new figures in the supplement (Figure S5 and Figure S6). We also have expanded the information provided in Figure 3. Finally, we have added in a new table (Table 4).

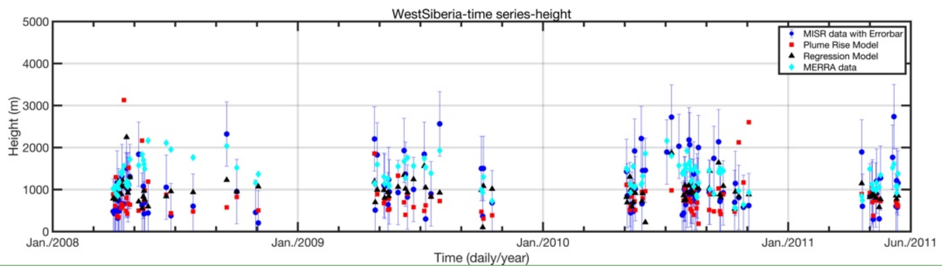
Changes made to the spelling and grammar are clearly shown in the track-changes version of the text itself.

The statistical analysis is unclear and in some cases contradictory and arbitrary (the authors describe predictors as orthogonal and then include a predictor that is a ratio of other predictors, data that agrees too poorly is thrown out).

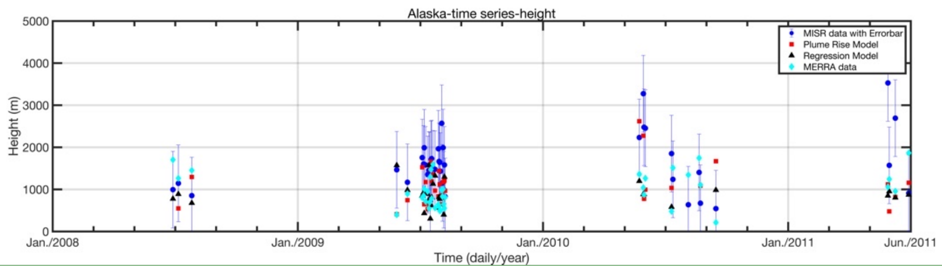
As discussed above, the ratio predictor has its own unique physical meaning, describing the ratio of the temperature to the amount of biomass burned at the instantaneous point and time where the burning occurred. The pure NO_2 term is also an instantaneous term, describing the temperature of the burning at the time of burning. This is consistent with the fact that the lifetime is NO_2 is very short, lasting far less than the day-to-day gap between the measurements. The pure CO term on the other hand is not an instantaneous term, instead describing the total amount of biomass burned over the past day (or days in the case of missing data) between the prior measurement and the most recent measurement. This result is also consistent with the long lifetime in-situ of CO, lasting from weeks to months, as described in Lin et al., 2020a, 2020b. Ideally for future work, we can find a third completely independent measurement which can also provide us a similar piece of knowledge such as provided by the term $[NO_2]/[CO]$, however such may not be possible until the next generation of satellite products is released to accomplish such a goal (i.e. Qin et al., 2020).

We agree in full about providing the full set of data over all of the areas. To better facilitate this, we have included more data in **Figure 3 and Figure S3**, including in regions where neither the egression model or plume rise model are found to be good fits. We also have included some extra discussion of these points. Furthermore, we have included an analysis of the black carbon height based on the mean daily MERRA hydrophobic black carbon values on the same days corresponding to where we have MISR height measurements. The data is provided in **Figure 3 (a)-(f)**.

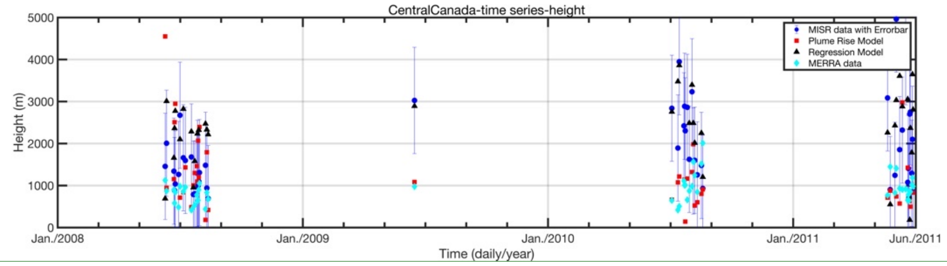
The following Figure has been added as Figure 3



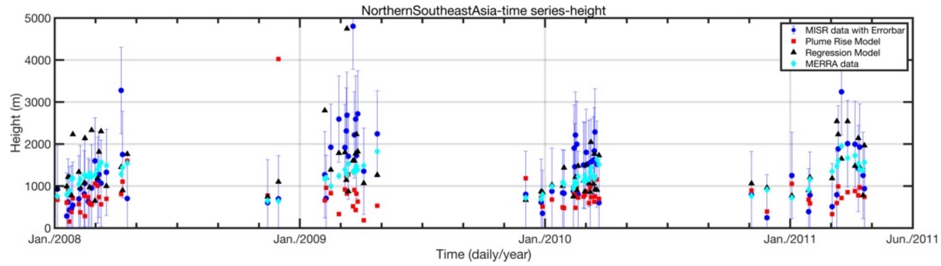
(a)



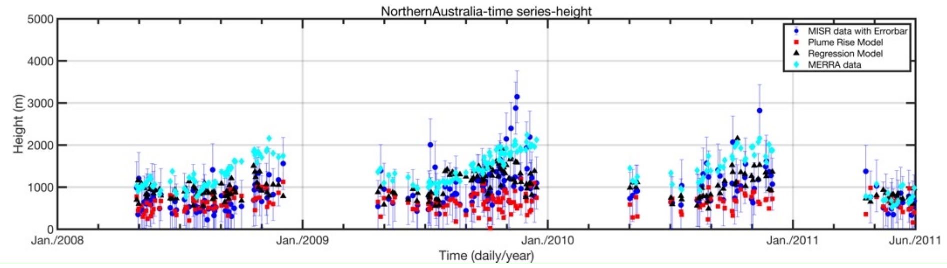
(b)



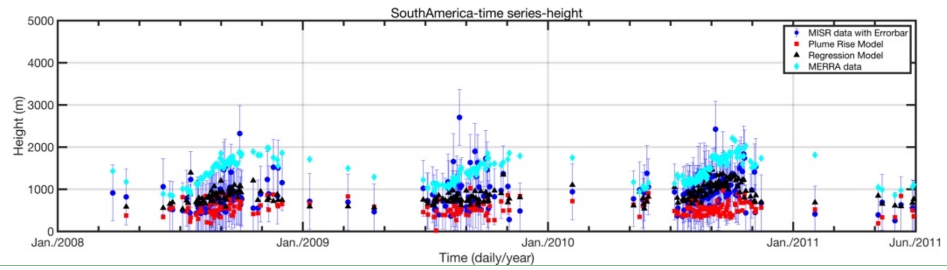
(c)



(d)



(e)



(f)

Figure 3: Time series of daily average measured MISR aerosol height (blue circles [m]) with an error bar corresponding to 1 sigma (blue bars [m]), the Plume Rise Model height (red squares [m]), the regression model height (black squares [m]), and the MERRA hydrophobic black carbon mean height (blue diamonds [m]). Part (a) corresponds to West Siberia, part (b) to Alaska, part (c) to Central Canada, part (d) to Northern Southeast Asia, part (e) to Northern Australia, and part (f) to South America. Missing data points are due to a lack of MISR measurements and/or measurements of regression model predictor(s).

Secondly, we have computed the statistics of the 10%, 30%, median, 70%, and 90% percentile heights of the daily MERRA hydrophobic black carbon heights on the same days where there are also MISR measurements. These results have been combined with the computed the statistics of the 10%, 30%, median, 70% and 90% percentile heights of the MISR measurements, the Plume Rise Model results, and the regression model results, into an expansion of **Table 3** and the new **Table 4**.

The following Paragraph has been added into the paper in Section 3.5

A comparison between the overall performance of the Plume Rise Model, the regression model, and MERRA leads to a few conclusions (Table 3). First of all, where the regression model exists, it reproduces the MISR height better than both the Plume Rise Model and MERRA. This includes over regions where the overall RMS error is very low such as Eastern Siberia and South America, as well as regions where the overall RMS error is large, such as Central Canada. This is true including over regions in the Arctic as well as in the tropics. Secondly, over the regions in which the regression model does not exist, MERRA provides a better reproduction of the MISR height than the Plume Rise Model in all cases, except for over Argentina. Perhaps this is true because of the fact that although MERRA uses data assimilation and a plume rise model type of code built in, the sharp height rise of the Andes Mountains and high cloud cover over this region lead to challenges that the global MERRA model cannot handle well. The second possible explanation is that the overall height of the plume is very low over Argentina and the local meteorology and FRP values are quite similar, which play to the plume rise model's strengths.

The following has been placed into the paper as Table 3

	MISR data	Plume Rise Model	RMS	Regression Model	RMS	MERRA Data	RMS
Central Africa	1.36 *(0.80)*	0.59 *(0.22)*	0.95	NAN	NAN	**1.72 *(0.50)***	**0.56**
Midwest Africa	0.90 *(0.42)*	0.60 *(0.23)*	0.47	NAN	NAN	**1.42 *(0.45)***	**0.41**
South Africa	1.71 *(0.56)*	0.58 *(0.23)*	1.18	NAN	NAN	**1.64 *(0.50)***	**0.44**
Central Siberia	1.64 *(0.90)*	0.87 *(0.89)*	1.01	NAN	NAN	**2.11 *(1.01)***	**0.66**
Siberia and North China	1.27 *(0.97)*	0.80*(0.64)*	0.69	**1.07 *(0.30)***	**0.42**	2.06 *(1.20)*	0.52
Eastern Siberia	1.12 *(1.00)*	0.68*(0.34)*	0.52	**1.32 *(0.65)***	**0.35**	3.13 *(1.09)*	0.68
West Siberia	0.95 *(0.77)*	0.79 *(0.95)*	0.67	**0.97 *(0.29)***	**0.47**	1.71 *(0.84)*	0.53
Northern Southeast Asia	1.57 *(1.03)*	0.73*(0.38)*	1.04	**1.42 *(0.51)***	**0.68**	1.40 *(0.63)*	0.75
Northern Australia	0.90 *(0.62)*	0.64*(0.29)*	0.57	**1.12 *(0.38)***	**0.52**	1.69 *(0.63)*	0.59
Alaska	1.57 *(0.91)*	1.39 *(3.03)*	0.88	**1.26 *(0.45)***	**0.77**	2.48 *(0.97)*	1.01
Central Canada	1.97 *(1.26)*	1.73 *(2.19)*	1.36	**2.13 *(1.72)***	**1.20**	2.54 *(1.17)*	1.36
South America	0.97 *(0.66)*	0.50*(0.21)*	0.52	**0.95 *(0.22)***	**0.37**	1.92 *(0.91)*	0.60
Argentina	0.69 *(0.70)*	**0.65 *(0.25)***	**0.40**	NAN	NAN	1.30 *(0.49)*	0.52
Eastern Europe	1.41 *(1.05)*	1.27 *(2.67)*	0.85	NAN	NAN	**1.15 *(0.59)***	**0.65**

Table 3: Statistics of measured MISR plume heights and (standard deviations) (2nd column [km]) using all available daily data from Jan 2008 to Jun 2011; Plume Rise Model heights and (standard deviations) (3rd column [km]); RMS error between the MISR plume heights and Plume Rise Model heights (4th column [km]); regression model heights and (standard deviations) (5th column [km]); RMS error between the MISR plume heights and regression model heights (6th column [km]); MERRA daily mean hydrophobic black carbon heights and (standard deviations) (7th column [km]); and finally the RMS error between the MISR plume heights and MERRA daily hydrophobic black carbon heights (8th column [km]). NaN indicates that the regression model failed over the respective region. The model type with the lowest RMS error over each region is given in "Bold".

The following Paragraph has been added into the paper in Section 3.5

Furthermore, comparing the performance of the plume rise model, the regression model, and MERRA at different percentiles of height leads to additional conclusions. On one hand, the regression model is the only one which does not have an obvious bias versus MISR measurements, with the regression model sometimes overapproximating and other times underapproximating different geographic locations at different height levels. In fact, the results at the median and 70% height levels are an excellent fit for 4 of the 8 different regions. On the other hand, both the plume rise model and MERRA have obvious biases. The plume rise model is almost always too low, with the only exception being its ability to model 6 of the 14 regions reasonably well at the 10% height level (i.e. the bottom of the plume). However, in the case where the 10% level is higher than other cases, such as a very narrow distribution, the plume rise model still dos a poor job. MERRA is almost always too high, with it performing best at only South Africa and

East Europe. Furthermore, the results from the plume rise model tend to also be narrower than the data, while the results from MERRA tend to be broader than the data. The results of MERRA being broad, as demonstrated clearly in Fig. 4, are not due to a high inter-annual variability, which actually barely exists in the MERRA dataset as compared with the regression model and MISR, but instead due to too much aerosol being found too high in the atmosphere, as well as too much aerosol being found at the surface.

The following has been added to the paper as Table 4

	MISR 10%	MISR 30%	MISR 50%	MISR 70%	MISR 90%	PRM 10%	PRM 30%	PRM 50%	PRM 70%	PRM 90%
Central Africa	0.70	0.99	1.22	1.53	2.10	0.33	0.47	0.57	0.68	0.85
Midwest Africa	0.43	0.69	0.87	1.05	1.37	0.30	0.49	0.60	0.70	0.85
South Africa	1.12	1.44	1.67	1.92	2.31	0.32	0.46	0.56	0.67	0.84
Central Siberia	0.75	1.15	1.48	1.93	2.62	0.38	0.59	0.74	0.91	1.27
Siberia and North China	0.58	0.92	1.15	1.41	1.88	0.38	0.55	0.68	0.84	1.24
East Siberia	0.41	0.77	1.00	1.29	1.69	0.36	0.49	0.62	0.78	0.97
West Siberia	0.28	0.56	0.79	1.09	1.71	0.38	0.52	0.62	0.76	1.14
Northern Southeast Asia	0.48	0.87	1.35	1.91	3.03	0.32	0.55	0.71	0.84	1.10
Northern Australia	0.28	0.56	0.79	1.09	1.52	0.34	0.49	0.63	0.75	0.93
Alaska	0.59	1.02	1.43	1.88	2.78	0.52	0.83	1.00	1.20	1.56
Central Canada	0.72	1.16	1.73	2.36	3.51	0.51	0.74	0.98	1.68	3.04
South America	0.38	0.64	0.85	1.11	1.65	0.26	0.39	0.50	0.60	0.77
Argentina	0.14	0.34	0.51	0.75	1.26	0.34	0.50	0.63	0.76	0.97
East Europe	0.44	0.85	1.19	1.60	2.63	0.47	0.64	0.82	1.08	1.97

(a)

	RM 10%	RM 30%	RM 50%	RM 70%	RM 90%	MERRA 10%	MERRA 30%	MERRA 50%	MERRA 70%	MERRA 90%
Central Africa	nan	nan	nan	nan	nan	1.08	1.47	1.71	1.96	2.33
Midwest Africa	nan	nan	nan	nan	nan	0.87	1.18	1.40	1.62	1.99
South Africa	nan	nan	nan	nan	nan	1.01	1.35	1.62	1.90	2.29
Central Siberia	nan	nan	nan	nan	nan	0.87	1.51	1.99	2.53	3.49
Siberia and North China	0.89	1.02	1.13	1.27	1.50	0.55	1.27	1.92	2.64	3.74
East Siberia	0.95	1.41	1.66	1.88	2.66	1.72	2.57	3.14	3.72	4.56
West Siberia	0.72	0.84	0.93	1.03	1.22	0.67	1.22	1.63	2.06	2.81
Northern Southeast Asia	0.81	1.00	1.20	1.69	2.64	0.68	0.99	1.29	1.65	2.29
Northern Australia	0.71	0.87	1.04	1.25	1.53	0.91	1.29	1.64	2.01	2.52
Alaska	0.30	0.80	0.82	0.85	1.35	1.25	1.94	2.43	2.94	3.76
Central Canada	0.80	2.01	2.28	2.78	4.59	1.02	1.81	2.49	3.22	4.13
South America	0.71	0.86	0.98	1.11	1.36	0.90	1.38	1.77	2.22	3.19
Argentina	nan	nan	nan	nan	nan	0.70	1.01	1.25	1.52	1.94
East Europe	nan	nan	nan	nan	nan	0.43	0.78	1.09	1.40	1.90

(b)

Table 4: Statistics of the 10%, 30%, median, 70% and 90% percentile heights [km] of MISR heights and plume rise model heights (a), and regression model heights and MERRA heights (b). NaN refers to regions where there is no regression model result.

The following paragraphs have been added to the paper in Section 3.5

The MISR data, regardless of the region, shows some amount of inter-annual variability. This ranges from a minimum over East Siberia and Siberia and North China, to a maximum over Central Canada and Northern Southeast Asia. On the other hand, MERRA shows only a very small variation anywhere, with most of the years exactly the same as each other. The amount at the surface is always much larger than found in MISR and the amount in the middle free troposphere is also much larger than in MISR. The largest variation in MERRA is found in Central Canada, Alaska, and Northern Australia. All of these are regions which are relatively cloud free and have vast amounts of ground stations, and therefore will have a large amount of the total MERRA model contribution from reanalysis data.

In the case of East Siberia there is only burning observed by MISR in 2 of the 4 years studied here, although these two different years have quite a different distribution. In 2008, the aerosol is limited in height to under 1000m, while in 2010, the aerosol has a peak height at 1000m and a significant fraction up to 2000m. In the case of Siberia and North China, the peak ranges from 800m to 1200m and the maximum ranges from 2200m to 3000m. MERRA shows no burning at all in East Siberia, with a completely flat profile all 4 years, and a consistent burning year to year, with the aerosol all confined to 1000m and below over Siberia and North China. In terms of the regression model, the fact that there is a good fit is supported by Fig. 5. As can be observed, all of the fire data points occur in regions of high CO and the vast majority also occur in regions of high NO_2. In Siberia and Northern China, the findings in both of the years in Fig. 5 lend support, albeit from two different perspectives. The first is that the fires always overlap with regions of high CO, and that in the 2011, one of the major differences is that the region in the middle has low CO and no fires, which were both present and highly polluted in 2008. The NO_2 is always high over the southern region, and is never very high in the central or northern regions, likely due to the intense cold air present in these regions altering the NO_2 chemistry.

Over Central Canada the MISR data shows peaks or sub-peaks at 1000m in 2008, 2800m and 3200m in 2009, 2000m in 2010, and 1000m and 2600m in 2011. In many of these years the amount located in the free troposphere is much larger than the amount in the boundary layer. Yet, even though this is the region in which MERRA has the most inter-annual variability, in all cases, the vast majority of the aerosol is found below 1000m. Furthermore, no peaks or subpeaks are found anywhere above the surface. Finally, MERRA only shows 1 year to be considerably different from the others, whereas the MISR data shows that all 4 years are quite different. By looking at Fig. 3, we can see that the regression model on some days underestimates the plume, on some days overestimates the plume, and on some days is nearly perfect. There is no bias, and the fact that it is able to capture the range of values over all 4 years indicates that the performance is not only better on average, but as well at capturing the inter-annual variation over this region. This finding is further supported by Fig. 5, where all of the MISR fire points in Central Canada in 2010 are found in high CO pixels, and most of the MISR fire points are also found in high NO_2 pixels. This demonstrates that the vast majority of the MISR plumes are local in nature and actively connected with the ground (due to the short lifetime of NO_2), are in relatively cloud-free regions where these remotely sensed platforms will work, but not necessarily MODIS which may be blocked by the high AOD levels, while also being in regions which are clearly heavily polluted by CO during these times, but are not normally so.

The MISR measurements over Northern Southeast Asia show the majority under 1000m but a second peak around 2500m in 2008, the peak at 2500m and a large amount up to 3200m in 2009, the peak was spread from 500m to 2500m in 2010, and peaks at 1000m, 1200m, and 2200m in 2011. This huge amount of inter-annual variability is not at all captured by MERRA, which is consistent with other recent findings over this area of the world demonstrating that many products based on MODIS tend to have problems (i.e. Cohen 2014, Cohen et al., 2018). However, the regression model performs well over this region as over all of the years, with measurements again showing an unbiased representation in all 4 years of the height, with some days high, other days low, and some days nearly perfect. This is in part demonstrated clearly in Fig. 3 and Fig. 5 by the fact that the MISR fire points occur over the highest loadings of CO and NO_2 found among any region, anywhere else in the world, as observed in this study.

The following Figure has been added as Figure 4

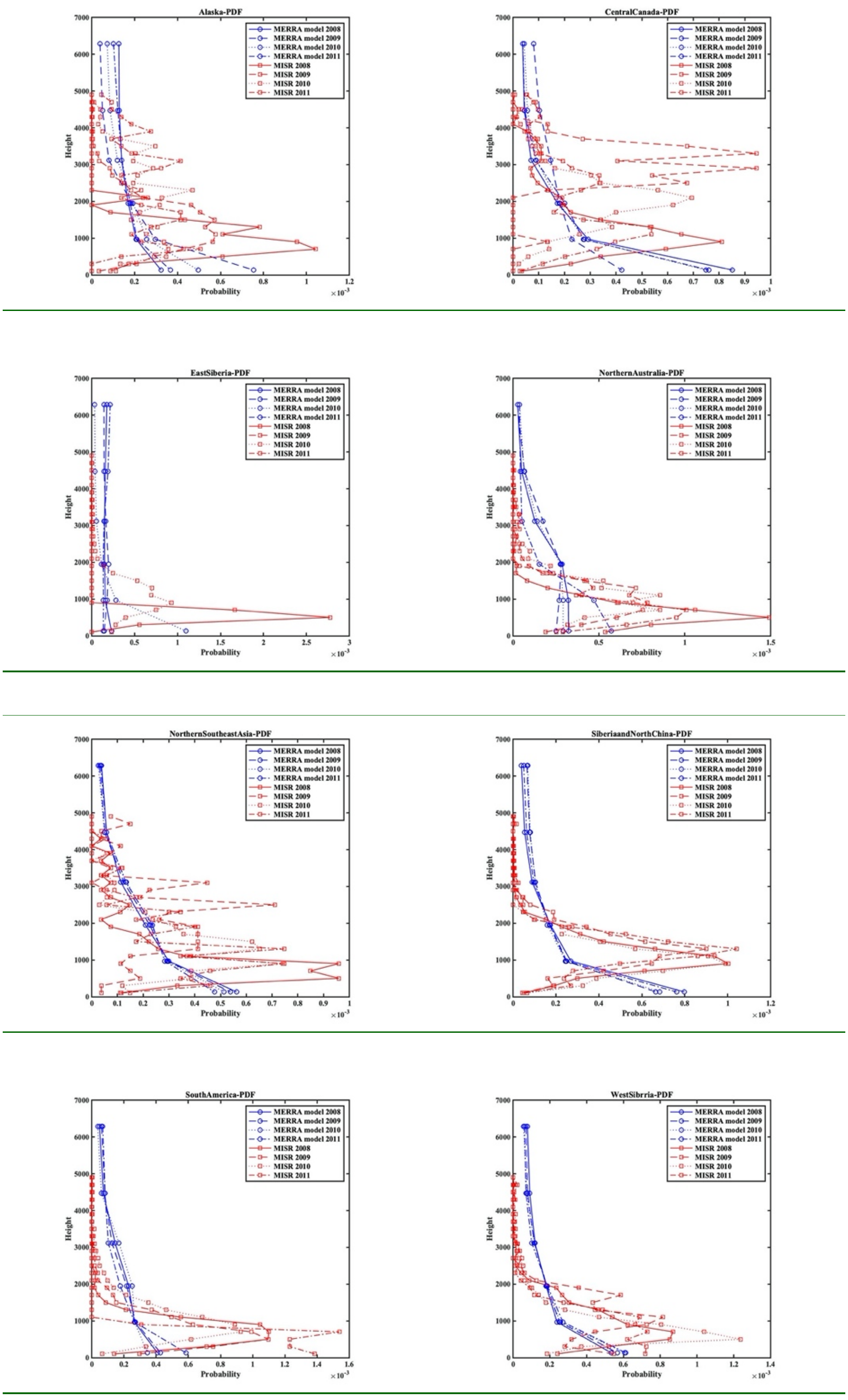


Figure 4: PDF of the vertical distribution of MISR heights (red lines for 2008, red dashes for 2009, red dots for 2010, and red dash-dots for 2011) and MERRA hydrophobic black carbon heights (blue lines, color scheme the same as for MISR). These plots are only

over regions in which the regression model applies.

There is no comparison of these results to any sort of reasonable chemical transport model (for instance MERRA2 might even have sufficient data to tell us about plume height and would be a fairer comparison).

In terms of the RMS error of the mean height over the entire time period, we determine that the MERRA model performs more poorly than the regression model at all places where the regression model passes the test of reliability. We also note that the MERRA RMS error is lower at the locations where the regression model does not pass the reliability test than over regions where it does the pass reliability test. This interesting result may further strengthen the idea that the regression model is accounting for some aspect of non-linearity which the underlying model used for MERRA is not accounting for.

MERRA performs better than the plume rise model in 8 regions, worse in 5 regions, and similarly in 1 region. Again, it is interesting to note that the region where the plume rise model works better than MERRA that does not also work for the regression model is in Argentina. Therefore, in general, these results show that the plume rise model almost never adds value, as compared to MERRA or the Regression approach, except for in Argentina. In the case of Argentina, MERRA has an obvious high bias, possibly due to the effect of the Andes Mountains being a dominant feature over much of this region's total area, and the known problems of global-scale models in representing highly mountainous regions.

Because I feel that the amount of work to add additional plume models, make the re- egression analysis more objective, and incorporate some chemical transport modelling results requires more work than can be accomplished in a review period I recommend rejection.

L18 Just saying the MLR model does a better job is a bit disingenuous. Linear least squares will always maximize variance explained. The authors need to show that they do some sort of out of sample testing.

We believe that the explanations above and comparisons with the Plume Rise Model and MERRA show that the MLR model does a better job. We understand clearly the concept of out of sample testing, but believe that it is not required in the case where, we are training against MISR and comparing against MERRA, that it is not required. We are not using the same dataset for training and comparison. Recall that as a data assimilation product, MERRA should be based on information which is quite different from the MISR plum heights, NO_2, and CO used in the training and comparisons.

The following sentences have been added to section 4

Our results show clearly that where we can successfully form a regression model, that it performs better than both the plume rise model and MERRA. The specific forms of the regression model that are the best are those which have NO2 or a combination of NO2 and CO (in particular when the non-linear term NO2/CO is considered). These results are consistent with our hypothesis and literature review that show new forms of non-linearity relating plume rise height to factors influencing buoyancy, radiative transfer, and energy transfer in-situ, and/or biases in remotely sensed measurements of FRP and land-surface products are important. Such are not considered in the present generation of plume rise models (including the global-scale models underlying MERRA). In the cases where we cannot form a regression model, we find that MERRA performs better than the plume rise model everywhere, except for Argentina, which has a unique high mountain just upwind in the Andes, coupled with a very low overall height, all of which are disadvantages for the models underlying MERRA. In general, this shows that improved model complexity and data assimilation doe produce a better result, as expected.

We propose the results as a first step of a new approach to parameterization that my help us to move forward in terms of improving our ability to reproduce heights of fire plumes for regional and global scale modeling and analysis studies over many different periods of time. We believe that our sample dataset is currently not sufficiently long to form an ideal fit, and hence thought that excluding data to self-compare was not an ideal use of the very limited resources we had. We do hope that as more new datasets are released, the community will have access to more relevant input data, and as more MISR plume height data is released, the community will have more access to better understand the vertical distribution of height.

L32 Use of significant should be reserved for statistical statements. Consider using ‘substantial’.

Thank you. This has also been implemented in other places as well.

L34 ‘and are known’

This sentence has been clarified.

L35 I believe biomass burning is also emitted at the surface and you mean it is moved into the upper atmosphere.

I would argue that the emission also does not occur at the surface, but instead occurs at wherever the material being combusted is in direct contact with the atmosphere, whether it is bubbles formed under the soil at the intersection of oxygen and peat, or it is in pieces of lofted grass which not yet fully burned but are caught in the uprising atmospheric plume and finally combust far above the surface.

The point that we all agree on is clear however: the emissions occur into parcels of air which rise at a sufficiently rapid rate that they are for all effective purposes of the measurements employed in this work (MISR, OMI, MOPITT, MERRA, and MODIS), “emitted” into the atmosphere at a given height. Sentence 35 has now been edited to reflect this.

L40 The statement that aerosols above the PBL have a bigger influence on the atmo- sphere may be true in some context, and the authors do provide citations, but they need to be a bit more specific here. I assume they mean in some sort of normal-ized sense (eg Pinatubo had a big influence on global mean temperature, but in an integrated sense aerosol in the boundary layer probably has a bigger impact). Either way, while a very interesting point to make, the authors might want to expand on this statement a bit for clarity.

This is the issue of radiative forcing. In this paper, we are looking at remotely sensed measurements on a scale of 1km to 100km, and hence at the implied radiative forcings at these scales. A very interesting topic for another time. The review paper included Tao et al. 2012 (already cited) is an excellent introduction to this topic.

L45 Who used? I think the authors have a typo and all the citations have stuck together.

Thank you.

L53 Lidar isn't capitalized: https://www-calipso.larc.nasa.gov/

Thank you for this correction. This has been implemented in 2 places.

L81 Large majority is redundant

Updated.

L99 typo, remove 'the'

Thank you.

L144 Specifically

Thank you.

L145 Does this mean that when you have cloud or aerosol you don't get CO measure- ments?

This is now clarified in detail based on a question from Reviewer #2.

L156 NO2 also has substantial industrial sources. The way that this is written implies that NO2 is only from fires.

We did not mean to imply that NO_2 does not have a significant urban source. We fully agree that NO_2 has a significant urban source. But we stated that the temporal-spatial distribution of urban NO2 is much lower than for fires, because other than transportation sources, most urban sources occur in fixed locations, and even transportation sources tend to follow fixed pathways (roads, shipping lines, air routes, etc.). This has been clarified.

L187 Note that inputs are not necessarily orthogonal, unless you pretreat inputs some- how. For example, NO2/CO is going to be correlated with NO2 and CO.

This has been explained above.

L188 Typo in this sentence.

Corrected.

L216 This sentence is very unclear- how are you 'injecting additional information'? As you say earlier all data sets have to be present. This seems to imply that data points with missing data will sometimes be considered and additional information will sometimes be 'injected'.

This sentence has been changed and broken into two.

L218 It is also unclear how you intend to reduce bias. Do you mean that you will try out data sets that measure the same quantity to get an estimate of bias.

See above.

L254 It would be good to define FRP somewhere in the intro or methods in terms of its physics (for people outside the biomass burning community).

The following has been added into section 2.7

FRP is the measure of the radiative energy released by the fire. It is usually found in the infrared part of the spectrum as this is the part of the EM spectrum that corresponds closely with the temperatures that fires occur at in the Earth System.

L270 something that I think needs to be discussed in the use of this plume rise model is that it is based on a model from 1965. In the methods there need to be a few sentences on why this model has not been improved upon since then, or why it is an appropriate comparison to the MLR model. Not discussing this runs the risk of making the plume model seem like a straw man to those outside the plume modelling community. Another aspect of this plume rise model is that earlier the authors state that it begins to fail for small fires. The analysis should really be subset to fires that satisfy the assumptions going into the model, rather than degrading the model with fires that the plume rise model is not designed for.

First off, the reason why the plume model fails for small fires is not because of an inherent problem with the plume rise model itself. It is with the fact that small fires are frequently missed altogether, or have their FRPs severely underestimated. This is not a problem with the plume rise model itself, but of the inputs being used inside of the plume rise model. Another issue is the resolution at which MERRA and most reanalysis meteorological products release their temperature and wind profiles, leading to too coarse of a resolution. You are right that a deeper analysis may be helpful. However, this was done in a previous paper we authored (Cohen et al., 2018) and we are not sure if copying and pasting that would be helpful here or not.

However, to more fully address this issue, many such corrections and additions have been made throughout the text, as outlined both above and below. Please let me know if you think that these changes are sufficient.

L329 A citation to a review article here might be helpful.

A review of this has been added, Gunturu et al., 2009.

L332 Different than each other? Do you mean when the plume model and the mea- surements? If this is the case this also seems fairly arbitrary to be testing the model and throwing out the results when they are poor.

All of the data for the plume rise model is now included, whether the region fits well or not. Therefore, this sentence is removed.

L340 Is this just a function of bias from the plume rise model treating fires that are smaller and thus don't satisfy assumptions in the model?

This is not true. There is no bias in terms of the plume rise model being able to handle smaller fires, the problem is that smaller fires tend to have their FRP and other remotely sensed characteristics biased, since they are too small as compared to the spatial and temporal assumptions underlying the fields being measured.

L341 how well the data what?

Multiple changes have been made to these paragraphs. The finding is that it is the higher rising fires which are not reproduced by the Plume Rise Model, which is the exact opposite of what the reviewer and the community have focused on in the past. Again, this supports the conclusions made here that it is in fact missing physical forces, some extreme form of underestimation of FRP for medium and large fires, or a combination of these factors that is driving these differences.

The following is the partially retained and partially edited paragraph in Section 3.2

Next, we look at the difference from day-to-day at each of the sites which has a mean value less than or equal to 0.25 km. Using these results, we find that the mean daily difference between the plume rise model and the MISR measurements as a whole show a large amount of variation, with a global average of 0.44 km, a maximum of 1.13 km (in West Siberia), and a minimum of 0.04 km (in Argentina). Across all of the different regions we find that the plume rise model underestimates the plume height. Furthermore, we find that the differences between the Plum Rise Model and MISR are not normally distributed, with higher values not being able to be reproduced under any conditions, strongly indicative of a bias, in that somehow the largest, hottest, or most radiatively active fires are those being not reproduced well by the Plume Rise Model. In addition to this, we compute the RMS error (Table 3) as a way of quantifying overall how well the model and MISR match. The RMS is found to be considerably larger than the difference of the means, indicating that a small number of extreme values are dominating the overall results, which were found to be 0.67 km, 0.88 km, 1.36 km, 0.40 km, and 0.85 km in the respective five areas.

L344 While I understand the attraction of minimizing the number of figures, but this article only has 3 in the main text. I feel that the PDFs of modeled and observed plume heights could be moved to the main text.

The PDFs of the observed plume heights, along with many other figures, have been moved into the main text. All of the underlying data, including plots in the supplemental information, are available at:

As included in the Code/Data availability statement:
https://doi.org/10.6084/m9.figshare.10252526.v1 and https://doi.org/10.6084/m9.figshare.12386135.v1

L365 How does the analysis account for times when the area is very crowded with burning? How does it tell where plumes actually originate from? Can a plume from another fire be mistagged or affect plumes from a nearby fire?

This is a fair point. We have relied on the MISR data as being able to distinguish the individual plumes. However, a fair argument was made by Cohen et al., (2018) that this question is actually not the right one to ask. In reality, if an instrument such as MISR cannot distinguish the plumes from each other, then effectively, as far as any modeling system will be able to capture, or the atmosphere will be able to feel, they are a single plume. This has been discussed at length in the paper cited above. The only case in which this would possibly matter is if there is a bias between the plume height at equilibrium locally and that of a plume cloud regionally. However, one could argue that if the fires are packed so tightly, that they should be measured as a group and not individually.

L375 A clear list of assumptions in the methods would be good. I as- sume there is more than one plume rise model in the literature (for example https://link.springer.com/article/10.1007/s10661-005-1611-y). The authors must show results from at least two leading plume rise models to show that the poor results of the 1965 model are not just due to poor construction of the model and limitations in what it can do (and applying the model outside of its assumed conditions).

We have read this interesting review article carefully and have found that it supports our conclusion. In fact, even the plume rise model we are employing was not discussed. In fact, the only ways they have discussed are using mesoscale models (similar to the vertical approach employed by Cohen and Prinn, 2011), global scale models (similar to the vertical approach employed by Cohen and Wang, 2014), and reanalysis products (similar to MERRA as employed here). We have included in results of measurement constrained studies using lidar as well, and found that such methods still fail, in that they are training models of the same type.

We have included the following sentence in section 1

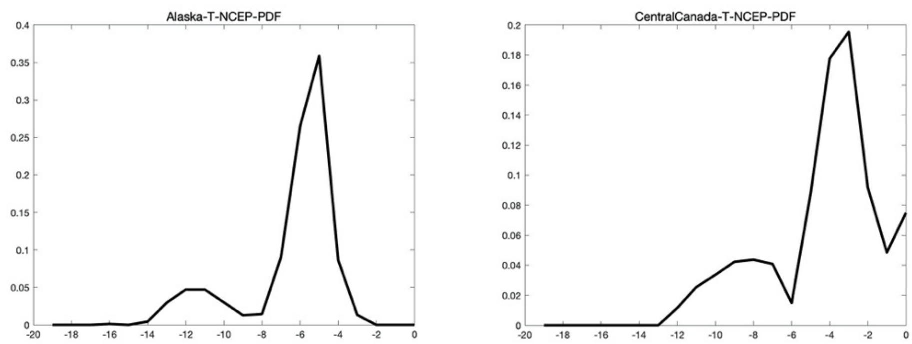
Large-scale reviews of the biomass burning literature spend a lot of time on how the atmosphere impacts the burning conditions, but also tend to overlook the issue of how the emissions are rapidly vertically distributed upon being emitted (Palacios-Orueta, et al. 2005).

L385 Is this because Argentina is dominated by the Pampas and fires tend to be over large areas and are uniform and the meteorology is relatively less complex?

Yes, this is also consistent with the results as demonstrated by Table 4, Figure 6, and Supplemental Figure S6.

We have edited and expanded upon the text to include the following sentence in Section 3.2

It is under these relatively lesser polluted conditions, where the fires are fewer and/or less intense, where a lower amount of total material is being burned on a per day basis of time over the total surface area burning, or where the meteorology and the vertical thermodynamic structure of the atmosphere are more uniform, that the plume rise model can achieve its best results (Table 4, Fig 6 and Fig S6), and thus that the plume rise model is reasonable to use in such a region.



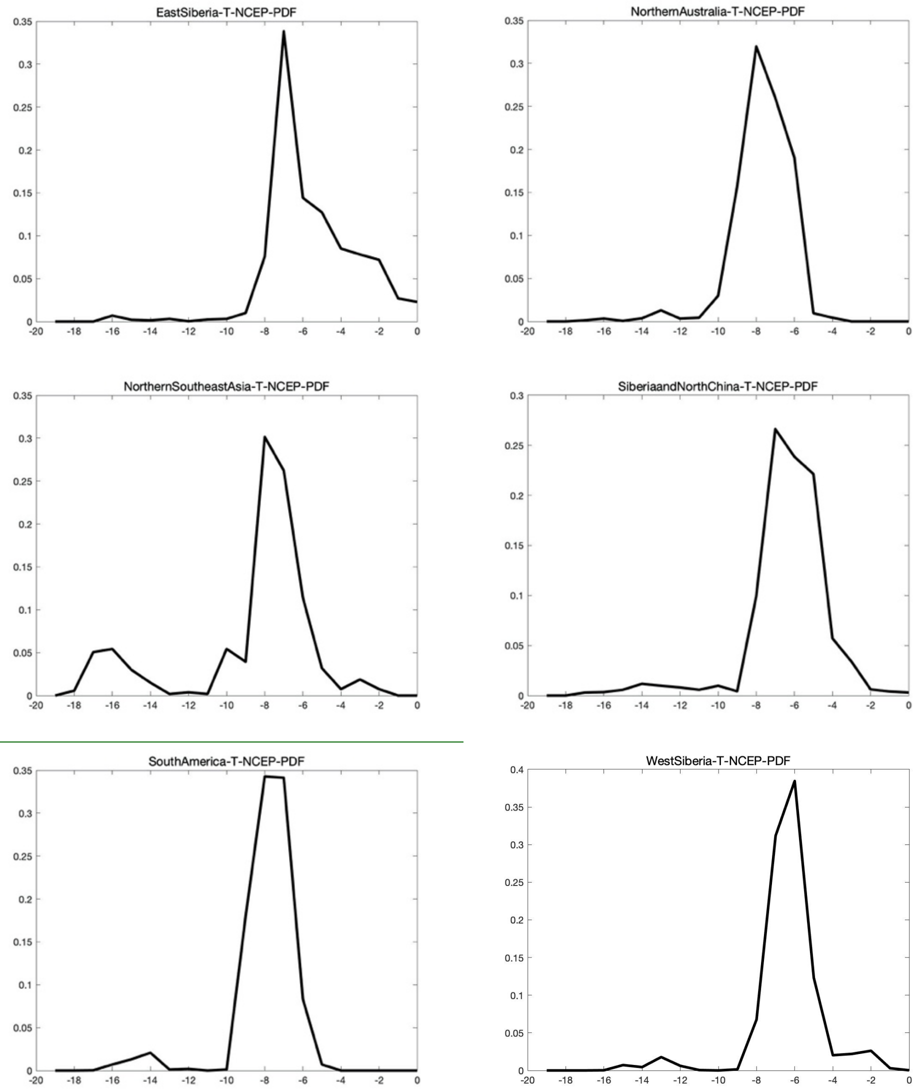


Figure 6: PDFs of the NCEP reanalysis vertical temperature gradient d[K]/d[km] over the locations and days that contain MISR plumes. The 8 regions over which the regression model is valid are shown.

L397 I think rather than coming up with 7 combinations of predictors a better approach might be to only have one model with all the predictors or use some sort of objective algorithm (eg machine learning) to remove low explained variance predictors. Arbitrar- ily coming up with 7 models seems like it will almost always guarantee a model works well.

Answered above.

L408 Fragment

Corrected.

L411 Again, I don’t understand how this is an evaluation if predictions that agree too poorly are removed.

Explained above.

L430 The three regions shown in Fig 3 are for a few plumes (judging by plotted data points) and for only a subset of the plumes in Fig1.

Figure 3 has been expanded.

L478 Which of the regression models is the new method?

Regression models 1 through 6 are new. The most useful are always regression models 1, 2, or 4. This has been explained in much more detail above.

L483 What are the 'modelled results' in contrast to the plume and regression models?

This has already been changed elsewhere.

L497 Somewhere there needs to a scatter plot of MLR model plume height versus observations. One possibility is that you are just fitting the mean. The MLR model is guaranteed to do this well (it minimizes unexplained variance). To do this correctly you should train the model on one region and apply it to other regions to get rid of the overfitting problem.

This has been explained above, and the results can be found in Figure 3, Figure 4, Figure 5, Figure 6, Table 4.

Fig1 I am not sure how useful this plot is because the dots obscure the land surface type.

We have updated Figure 1 with this

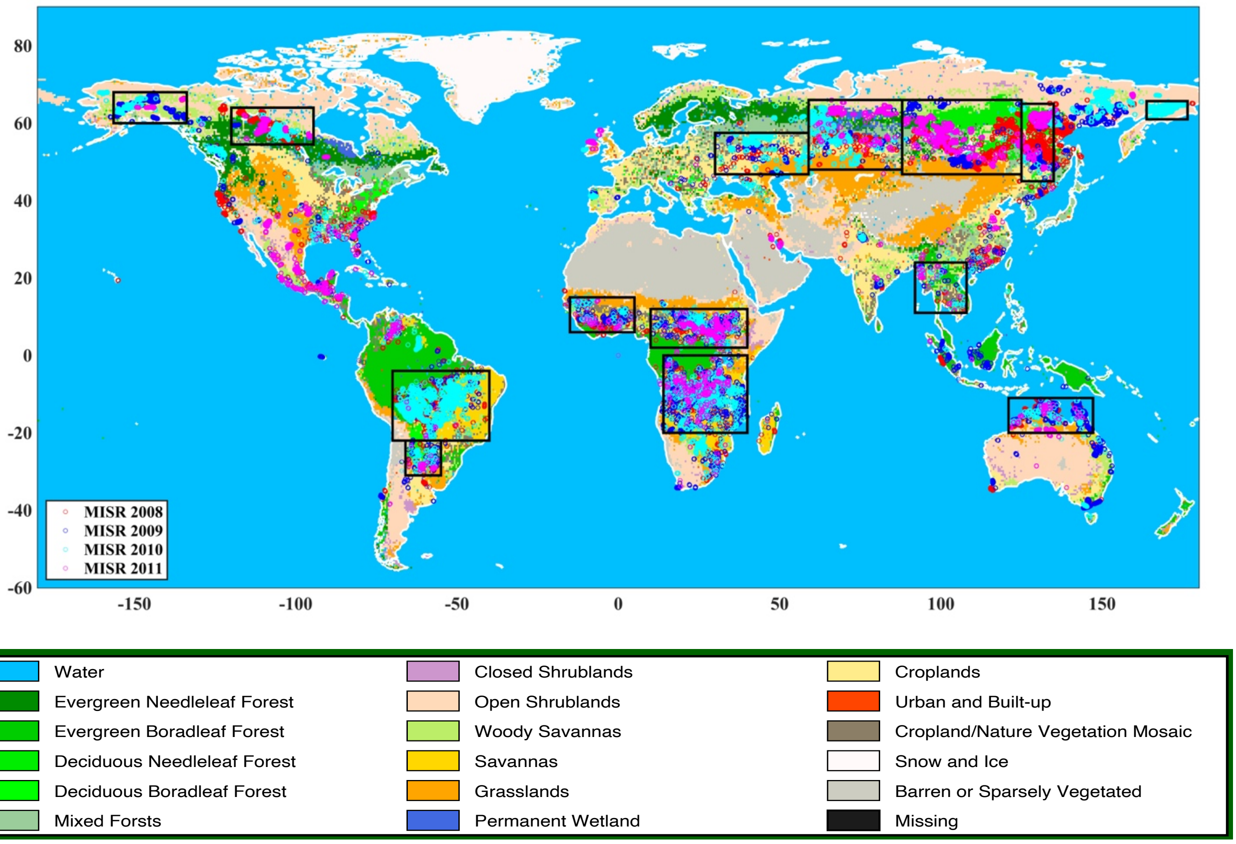
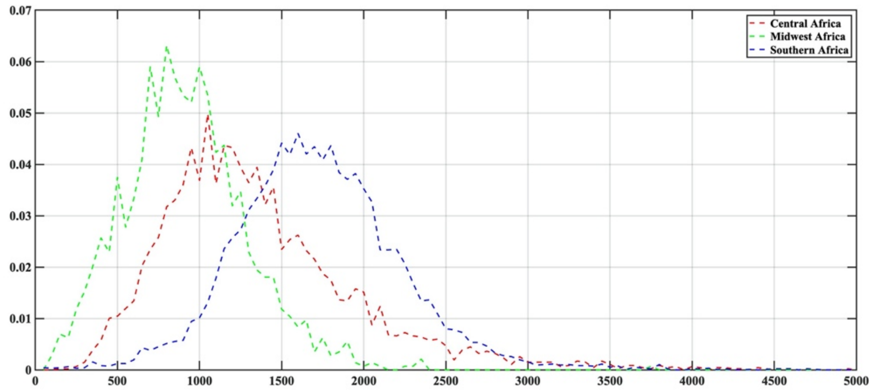


Figure 1: Land surface type at each of the daily MISR measurements from January 2008 to June 2011. Each dot corresponds to an individual aerosol plume, with different colors representing different years.

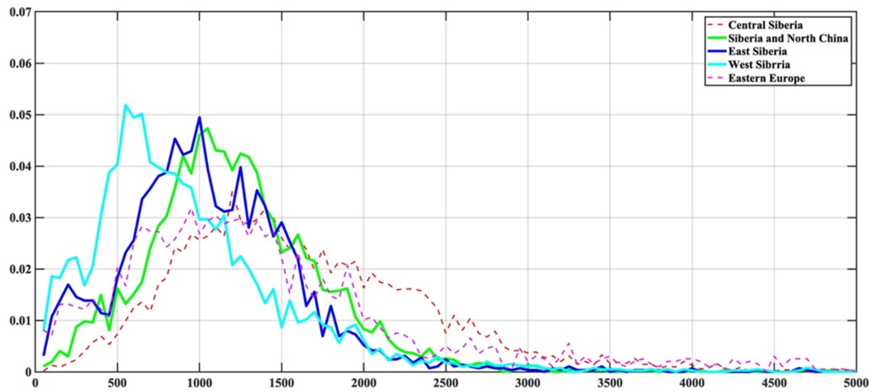
Fig2 Please use some different line styles and markers. Most of these colors are indistinguishable.

An excellent comment. We have made some changes here.

The following are now used for Figure 2

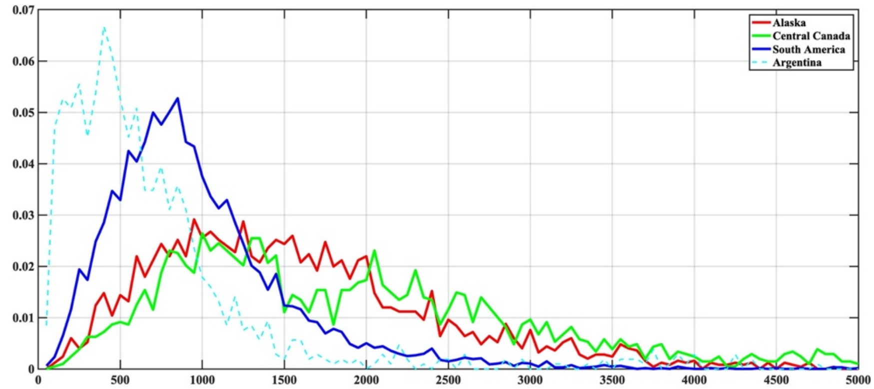


(a)



(b)

(c)



(d)

Figure 2: PDFs of all daily MISR plume height measurements from January 2008 through June 2011 (which are 5000m or less) over each of the following geographic regions: (a) Africa, (b) Eurasian High Latitudes, (c) Tropical Asia, and (d) the Americas. Solid lines correspond to regions which have a successful regression model, while dashed lines are regions which do not.

Response to Author 2:

I will keep this short and to the point. I think the basic idea of trying to investigate the relationship between trace gas/aerosol plume height and the pollutant loading is good. But having read the manuscript few times, I do not believe the authors have approached the problem with the right tools. My opinion/review is mostly from the observational perspective and I don't know much about the plume models.

1) Why use the total column values of NO2 and CO, when the authors themselves C1
show how, depending on the region, aerosols can be lifted to different heights. What do we actually scientifically gain by looking at the total column only? It is not a surprise that when episodes of strong pollution occur (e.g. fires, biomass burning), the total column values will increase and depending on the thermodynamical conditions (e.g. strength of convection) the lofting will occur. I understand that the vertically resolved observations of NO2 are not available, but altitude-resolved CO retrievals are available from a number of sensors, MOPITT, AIRS, IASI etc. I also wonder why the authors don't use aerosol layer heights from CALIPSO (possibly combined with OMI)? Wouldn't that be the most accurate account of plume heights?

We have introduced the results from MERRA into the paper, based on comments from the first reviewer. We do agree that additional vertical measurements from MOPITT would be interesting to investigate as well, but instead propose this for a future effort. One reason for this is that the horizontal and vertical resolution of MOPITT are very challenging to use unless very carefully applied, which would go beyond the current time allotted for this major revision. Furthermore, we are using actual measurements of height from MISR, and first wanted to see if the simpler column loadings would be representative. This further is consistent with recent findings from my team as just published in Lin et al., 2020a. which have shown that the column loading of highly variable regions of CO map very well with biomass burning events, as well as Lin et al., (Under Revision) 2020b; and Cohen et al., 2018, in which we have further looked into the MOPITT vertical distribution associated with global biomass burning, although at temporal scales of a week to months, not day-to-day as we are working on here. The suggestion of using CALIOP is also very interesting, and we believe that based on the results from Cohen et al., 2018, it would yield significant findings. Again, we did not have enough time in this current revision round to accomplish this, especially since finding a sufficiently large number of overpasses in a region which is actually influenced by the plumes, not merely over an "average region" is incredibly challenging work. However, we appreciate this suggestion and will seriously look forward in the future to address this.

As per your suggestion, we have carefully checked the NCEP vertical temperature gradient as a proxy of the thermodynamic conditions (e.g. strength of convection) and the vertical air mass rise at the surface.

Based on these findings, we have added the following to the paper in Section 3.5 and Figure 6

In terms of the magnitudes of the vertical temperature gradient (dT/dz) and the vertical wind speed at the surface, we have not found any correlation or relationship between the cases in which the regression model performs better or worse. Even considering those cases in which there are extremely atypical values in these variables, such as positive temperature gradients (i.e. an unstable atmosphere), or negative temperature gradients which are more negative than the -9.8 K/km rate which is the pure dry air thermodynamic limit (i.e. extreme stabilization due to intense aerosol/cloud cooling), as observed in Fig. 6. This provides a further piece of support to the idea that the regression model works well under conditions where there is some local non-linear forcing in the system which is not being taken into account, whether it is a coupled chemical, aerosol dynamical/size, radiative-dynamic, thermodynamic, or direct/semi-direct/indirect type of aerosol effect, all of which are being accounted for to some degree by the loadings of NO_2 and CO, but which are missed by the model underlying the meteorological reanalysis data (e.g. Cohen et al., 2011; Wang et al., 2009).

However, it does seem that under the conditions where the regression model was not able to be formed, that there are some important differences in terms specifically of the vertical temperature gradient variable. In specific, in the cases in which the value of dT/dz is either more negative than -9 K/km or positive, that the MERRA results are far better than those from the plume rise model, as compared to not under those conditions. However, such cases only account for 15% or fewer of the total cases observed in this study, and therefore do not play an outsized role.

2) The lifetimes of CO and NO2 are very different. CO has much more homogenized distribution in the atmosphere, especially as the altitude increases due to transport processes etc. So can the authors disentangle this background signal from the one that is associated with the biomass burning plumes for CO, especially over those regions that already have strong background variability in industrial+traffic pollution?

This is a great point and one of the reasons why we also wanted to choose to use both NO_2 and CO simultaneously.

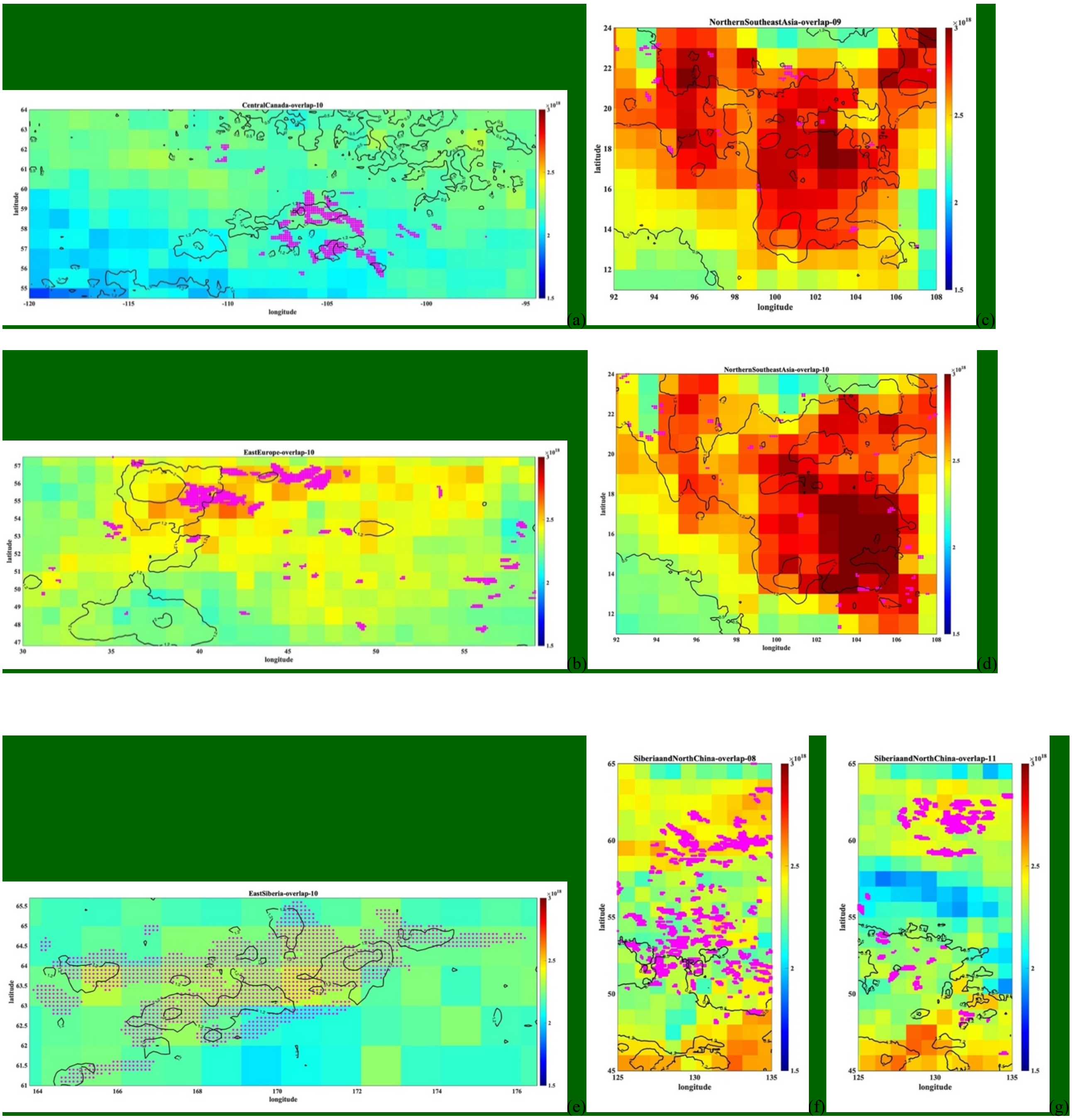
We have included the following text in section 3.3, Figure 5, and Table 1

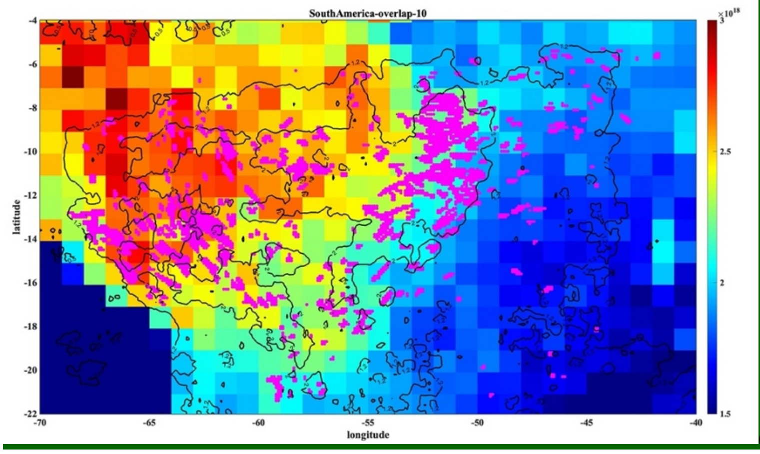
Due to the fact that NO2 and CO have very different lifetimes in the atmosphere, a fire-based source is expected to have a high level of both CO and NO2 close to its source, which decays as one heads away in space from the source. This decay should be a function of the wind direction as well, as both the CO and NO2 upwind will not have a significant source, but downwind the CO will have a significant source, as shown in Fig. 5. We find that our results are consistent with this theory. First, we have found that the regions that have the highest NO_2 at the same time as the MISR measurements are made, also have a very strong overlap well with the locations of the MISR plume heights. We further determine this to be true for every year on a year-by-year basis (Fig S1). Second, we find that the higher values of CO match well with the year-to-year locations of MISR fires (or downwind thereof) at most of the sites, including in Alaska, Central Canada, Central Siberia, East Europe, East Siberia, Northern Southeast Asia, Siberia and North China, and South America. As expected, there greater smearing away from the source regions. As expected, this is due to the fact that the lifetime of CO is much greater.

Furthermore, in terms of changes in time, a climatology of CO should be slightly higher due to the added emissions from the fires, but the NO2 should be much larger than the climatology, since there is little to no retention in the air, as demonstrated in Table 1. To account for this, we have also looked at the difference between the fire times and the long-term climatology. Over regions which are urban and hence contributing randomly to the variance, we expect the differences to be smaller than due to the fires, and this is observed clearly as well. These results are also shown to be consistent with recent work (Cohen, 2014; Lin et al., 2014; Lin et al., 2020a), showing that the characteristics of the spatial-temporal variability of fires is quite different from that of urban areas, and has a much higher variability both week-to-week and inter-annually.

Thirdly, this is pointed out in the time series plots **(Figure S1)**, where the CO and NO_2 are both considerably higher during the fire times than the rest of the year, while at the same time, the NO_2 and CO are both higher over the subset of points that have fires on the fire days than over the entire region on the fire only days. The idea of a proper study to disentangle the downwind regions from fires, downwind regions contaminated by both urban regions and fires, and downwind regions from urban-only regions is something of merit and would be an excellent follow-up work. This part of the response will not go into the main paper.

The following is now included as Figure 5





(h)

Figure 5: Spatial distribution of annual compilation of all MISR fires (magenta dots), mean OMI NO_2 column loading on days where there are fires (black isopleths [$*10^{15}$ mol/cm^2]), and mean MOPITT CO column loading on days where there are fires (Colorbar, mol/cm^2). The corresponding regions are: (a) 2010 Central Canada, (b) 2010 East Europe, (c) 2009 and (d) 2010 Northern Southeast Asia, (e) 2010 East Siberia, (f) 2008 and (g) 2011 Siberia and Northern China, and (h) 2010 South America.

3) There is virtually no description of how different satellite data products are quality controlled, analysed etc. The devil is in the details. What quality flags are used? How are cloudy/non-cloudy cases handled? Is there a consistency in such cases across all datasets? How is the sampling affected by the quality control?

To add in more details, the following have been added at the respective parts of the manuscript in sections 2.3, 2.4, and 2.7. Additional corrections have been made in 2.3 to reflect the updated version of the CO data used.

The following has been added to section 2.3

In terms of the CO from MOPITT, we take the day time only retrievals (to reduce bias) from version 8, level 3 data. In specific we use the combined thermal and near infrared product (Deeter et al, 2017). We further constrain the data to where the cloud fraction is less than 0.3 and where the vertical degrees of freedom are larger than 1.5. This combination has been shown to allow us to trust that there is a sufficient amount of signal and knowledge to demonstrate an actual measurement in the vertical, as compared with a result only dependent on the a priori model, as shown in Lin et al. (2020a). There are further gaps in the data due to orbital locations and very high aerosol conditions, all of which prevent entire coverage of our areas of interest each day. Therefore, we average all of the individual MOPITT data that passes our test to a 1°x1° grid.

The following has been added to section 2.4

In terms of the NO_2 from OMI, we first take the daily retrievals under the conditions where the cloud fraction is less than 0.3. Next, we aggregate the data to 0.25° x 0.25° using a linear interpolation and area weighted approach. In this way, those measurements near the edge of the swath or which are adjacent to cloudy areas are weighted less heavily in terms of the merged product. However, the areas are sufficiently large as to be roughly representative of the emissions from biomass burning of the NO_2 from within the grid box, as compared to that transferred from adjacent grid boxes.

Furthermore, for our computations we only retain those measurements in which we have data at the place of interest from MOPITT, OMI, and MISR at the same time. If just one of the three measurement platforms is more than 30% cloud covered, is not able to measure due to extremely high AOD levels, or is found outside of the swath at the given time, then that day's data is discarded in terms of developing and the regression model, and any subsequent analysis. However, we do use all available data every day from within the respective boxes in terms of understanding the background values, and trying to better constrain the differences between the values of the column measurements over the identified biomass burning points based on MISR and those which are within the same larger area but are upwind, downwind, or not involved with burning at all. This is completely consistent with the fact that biomass burning is sub-grid within each individual respective 1° x 1° box for CO and 0.25° x 0.25° box for NO2, while simultaneously only occurring over a distinct of set of days.

References:

Cohen, J. B. and Prinn, R. G.: Development of a fast, urban chemistry metamodel for inclusion in global models, Atmos. Chem. Phys., 11, 7629–7656, doi:10.5194/acp-11- 7629-2011, 2011.

Cohen, J. B., Hui Loong Ng, D., Lim, A. W. L., and Xin R. C. J.: Vertical distribution of aerosols over the Maritime Continent during El Niño, Atmos. Chem. Phys., 18, 7095-7108, doi: 10.5194/acp-18-7095-2018, 2018.

Deeter, M.N., Edwards, D.P., Francis, G.L., Gille, J.C., Martinez-Alonso, S., Worden, H.M., Sweeney, C., 2017. A climate-scale satellite record for carbon monoxide: the MO- PITT version 7 product. Atmos. Meas. Tech., 1–34. doi:10.5194/ amt-2017-71, 2017.

Ekman, A. M. L., Engström, A., Söderberg, A.: Impact of two-way aerosol-cloud interaction and changes in aerosol size distribution on simulated aerosol-induced deep convective cloud sensitivity. Journal of the Atmospheric Sciences, 68, 685-698, 2011.

Guo, J., Li, Y., Cohen, J. B., Li, J., Chen, D., and Xu, H.: Shift in the temporal trend of boundary layer height in China using long-term (1979–2016) radiosonde data. Geophysical Research Letters, 46, 6080– 6089. https://doi.org/10.1029/2019GL082666, 2019.

Ichoku, C., Giglio, L., Wooster, M., and Remer, L.: Global characterization of biomass-burning patterns using satellite measurements of fire radiative energy, Remote Sens. Environ., 112, 2950–2962, 2008.

Kaiser, J.W., Heil,A., Andreae, M.O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J., Razinger, M., Schultz, M.G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527–554, doi:10.5194/bg-9-527-2012, 2012.

Lin, C. Y., Cohen, J. B., Wang, S., and Lan, R. Y.: Application of a combined standard deviation and mean based approach to MOPITT CO column data, and resulting improved representation of biomass burning and urban air pollution sources, Remote Sens. Environ., 241:11720, https://doi.org/10.1016/j.rse.2020.111720, 2020.

Lin, C.Y., Cohen, J.B., Wang, S., Lan, R.Y., Deng, W.Z., Zhang, Y.L., and Liang, H.F.: New Approach to Understanding the Vertical Distribution of Biomass Burning Carbon Monoxide Based on Climatological MOPITT Vertical Measurements, (UNDER REVIEW, 2019).

Lin, N. H., Sayer, A. M., Wang, S. H., Loftus, A. M., Hsiao, T. C., Sheu, G. R., and Chantara, S.: Interactions between biomass- burning aerosols and clouds over Southeast Asia: Current status, challenges, and perspectives, Environ. Pollut., 195, 292–307, 2014.

Leung, F. Y. T., Logan, J. A., Park, R., Hyer, E., Kasischke, E., Streets, D., and Yurganov, L.: Impacts of enhanced biomass burning in the boreal forests in 1998 on tropospheric chemistry and the sensitivity of model results to the injection height to emissions, J. Geophys. Res., 112, D10313, https://doi.org/10.1029/2006JD008132, 2007.

Mims S.R., Kahn R.A., and Moroney C.M.: MISR Stereo Heights of Grassland Fire

Smoke Plumes in Australia. IEEE Transactions on Geoence and Remote Sensing , 48(1):25-35, 2010.

Qin, K., X. Han, D. Li, J. Xu, D. Loyola, Y. Xue, X. Zhou, D. Li, K. Zhang and L. Yuan.: Satellite-based estimation of surface NO2 concentrations over east-central China: A comparison of POMINO and OMNO2d data. Atmospheric Environment, 224: 117322, 2020

Tao, W.-K., Chen, J.-P., Li, Z., Wang, C., and Zhang, C.: Impact of aerosols on convective clouds and precipitation, Rev. Geophys., 50, RG2001, doi:10.1029/2011RG000369, 2012.

Wang, C., Kim, D., Ekman, A.M.L., Barth, M.C., and Rasch, P.J.: Impact of anthropogenic aerosols on Indian summer monsoon. Geophysical Research Letters, 36(21), 2009.

Zhu, L., Martin, M.V., and Hecobian, A.: Development and implementation of a new biomass burning emissions injection height scheme for the GEOSChem model. Geoscientific Model Development Discussions, 1-30, 2018.