Response to Anonymous Referee #1

Comment: This study represents a credible attempt at a new way to infer surface PM2.5 levels from CALIOP data, on a regional, two-year average basis. An advantage of CALIOP over passive sensors for this sort of analysis is the fact that it measures vertical profiles of backscatter and depolarisation, so bypasses a limitation inherent with imager data in partitioning between total column and near-surface aerosol loadings. In contrast, an acknowledged limitation is the curtain sampling of CALIOP vs. the broad-swath sampling of MODIS, etc. The authors introduce their technique and explain the relevant assumptions, and show results over the USA, evaluated with EPA monitors. This is a sensible, strong first step in this direction. The topic is important and relevant to AMT. I have a number of comments (below) but on the whole recommend that the paper can be accepted after minor revisions. Hopefully this will be a springboard for further studies refining the technique and expanding to other regions and time periods.

Response: We thank the reviewer for his/her comments and encouragement.

Comment: As a general comment, much of the quantitative evaluation is presented as scatter plots with linear regression fits, and the discussion is often framed in terms of r2 and slope. I'm not sure that this is the right thing to do here. One reason is that my understanding is that there can be non-negligible uncertainties on the PM data. Indeed, Ayers (2001, https://www.sciencedirect.com/science/article/pii/S1352231000005276) recommends using reduced major axis (RMA) regression instead of ordinary least squares when comparing PM monitors, for that reason. But also, the analysis in section 3 indicates that the CALIOP-derived estimates seem to have PM-dependence on their uncertainties too, so standard RMA may not be right either (as that assumes independent identically-distributed errors). For this reason I'd recommend Deming regression as a reasonable alternative (https://en.wikipedia.org/wiki/Deming_regression) when trying to compute the best-fit line. This should be more appropriate for this case, has packages in standard programming languages (and is not hard to code anyway), and is not hard to interpret. So this should be a pretty straightforward change to make which would improve the rigor of the manuscript. I recommend this is done throughout. Or, alternatively, don't fit a line but report something like mean ratio and RMS across certain ranges by binning the data.

I think it is important that appropriate statistical methods be used; continued publication using techniques we know to be deficient for our analyses just normalizes and encourages bad practice in the future. There isn't really a good justification for not fixing this.

Response: Thank you for the comments and suggestions. As recommended, Deming regression best-fit lines were added to the scatterplots of Figs. 1, 3, 4, 8, and 9, and the slopes computed from the Deming regression analyses were added to Tables 2, 3, and 4. Corresponding changes in regard to these figures and tables were made to the narrative, and the following was added to the end of Section 2 to describe Deming regression: "Lastly, we note that most of the results are shown in the form of scatter plots with fits from Deming regression (Deming, 1943). Due to uncertainties in $PM_{2.5}$ data, we show slopes computed from Deming regression analyses instead of those from simple linear regression. Deming regression in particular is appropriate here, as it accounts for

errors in both the independent and dependent variables (Deming, 1943), and has been used in past $PM_{2.5}$ related studies (e.g., Huang et al., 2014)."

Comment: My remaining comments are given as PXX, LYYY referring to page and line numbers respectively.

P1L21: I suggest replacing "sizes" with "diameters", as that is my understanding of the definition, but the remote sensing community often refers to radius instead when discussing size.

Response: Thank you for this suggestion. We have made the recommended changes.

Comment: P4L95: I am curious as to why, with over 10 years of data, the two-year period 2008-2009 is used here? If sampling is a limiting factor in some areas, surely adding a few more years would help with this? Is there something special about these two years, or some a priori reason why two years provides sufficient sampling? I realize that running the whole mission is probably not feasible at this stage. But I would imaging that in the time between this comment being posted and the close of Open Discussion, there wouid be sufficient time to download and analyze an additional few years of data. This should mostly be a matter of storage and CPU time, since the code is already written (and since the first author is at Langley where CALIPSO is based, I doubt computational concerns would be significant here).

Response: The two-year period of 2008-2009 was chosen because we wanted to be consistent with the temporal domain of our previous $PM_{2.5}$ study (Toth et al., 2014). An explanation is included in Section 2. We agree that adding more years would increase sampling, but we feel this is more appropriate for a future paper, as the purpose of this manuscript is to provide an initial demonstration of the concept. An extended analysis is planned for a forthcoming paper.

Comment: P6L123: somewhere in this initial paragraph, I'd ideally like some more discussion of the EPA data. For example, what are the uncertainties, is there any significant difference in these between the TEOM and BAM methods, and is there a difference in the siting of these two instrument types? If they're super-accurate and precise and equivalent, that's important to know. But if one is better than the other, and there's some spatial/temporal clustering in when TEOM vs. BAM is employed, that is also important to know. Recently, Kiss et al (2017, https://www.atmos-meas-tech.net/10/2477/2017/) published an analysis showing biases in hourly PM10 measurements. Is that relevant here? It might be, especially since that some daily averages in the EPA data correspond to a single sample. These are examples of things I'd like to see covered in the opening part of this section.

Response: Thank you for the comment. As for the Kiss et al. (2017) study, PM data with a lower temporal resolution (like 24-hour, "daily" data) are less biased compared to hourly data. Still, uncertainties in hourly data are likely to impact daily data that are averaged from hourly data. To fully quantify this issue would deserve a paper of its own. Here, as suggested by the reviewer, we have edited the discussion in this section to incorporate uncertainties of the various PM2.5 measurements and spatial representativeness of the different instruments/methods. The following was added to the text:

"Note that uncertainties have been reported for hourly PM measurements (Kiss et al., 2017). Examples of some uncertainties in these $PM_{2.5}$ measurements depend upon the instrument/method used: gravimetric (e.g., transport to the lab/human error and volatization of PM during the drying process; Patashnick et al., 2001), TEOM (e.g., errors due to improper inlet tube temperature; Eatough et al., 2003), and beta attenuation monitors (e.g., changes in the sample flow rate due to variations in temperature and moisture; Spagnolo, 1989). Also, it has been found that beta attenuation monitors may be more accurate than TEOM, as TEOM can underestimate $PM_{2.5}$ at low temperatures (e.g., Chung et al., 2001). Still, as suggested by Kiss et al. (2017), PM data collected over a longer period of time are much less likely to be biased. Thus, we expect lower uncertainties from data collected over 24-hours, then daily data generated by averaging hourly observations. Fully quantifying the differences from the two different PM observing methods, however, is a subject for a future study."

Comment: P8L189-190: This assumption (negligible mass above 10 micron size) is probably reasonable. But it would be fairly easy to try and quantify with AERONET. Take the inversion product from a half-dozen AERONET sites and count the fraction of the volume size distribution above 10 microns (and note here that the AERONET retrievals report size in terms of radius, while PM definitions are in diameter). You have to make some assumption about the density of particles being the same across the size range, but otherwise that gives a first order estimate at how big the effect might be, which could be compared to the other parts of the uncertainty analysis in section 3.2. I think AERONET dust radius peaks somewhere like 2.5 microns so in the western US, it might be that there's some dust contribution from the tail of the distribution which is being systematically missed here and would lead to an overestimate in the CALIOP-derived PM levels. Maybe it is negligible, but it would be fairly easy to show that it is negligible, and the authors have not.

Response: It is a nice idea but we think it might be difficult to apply the proposed idea for the US for a few reasons. Firstly, reliable AERONET volume size distributions are obtained from inversions that are performed when the 440 nm AOD is larger than 0.4 (Dubovik et al., 2006). In this study, we emphasize studying 2-year means over the US, which rarely exhibit averaged 440 nm AODs larger than 0.4. Secondly, we are only concerned with near surface (100-1000 m) aerosols for this study, but AERONET would provide values for the entire column, making such a comparison difficult. We argue that our assumption of negligible mass above 10 microns is reasonable because dust has been excluded from the analysis, and sea salt represents a small fraction of aerosols in the 100-1000 m atmospheric layer over the US for the 2008-2009 time period (i.e., $< 2\%$). Thus, we did not implement this change as suggested.

Comment: P13L299: An alternative to this (whether for the sensitivity analysis or the analysis as a whole) might be to look at the whole boundary layer (determining on a case by case basis) rather than testing different height ranges. Assuming that boundary layer depth is included as part of the MERRA2 meteorology being used here? This would go from assuming "the surface level of PM is represented well by the atmospheric layer from 0.1-1 km" to assuming "the boundary layer is well-mixed so represents the surface PM well", which is subtly different and might work better. I do agree that it seems reasonable to exclude the lowest 100 m, though.

Response: Thank you for this suggestion. Unfortunately, boundary layer depth is not included in the MERRA-2 meteorological profiles used for this analysis. MERRA-2 relative humidity was chosen for the paper because it was already collocated with the CALIOP aerosol profiles. A boundary layer depth analysis would not be a straightforward task, and we believe a thorough study into this important topic is best left for another paper during which our method can be further refined.

Comment: P14L323: This section made me wonder why the authors do not estimate PM10 from CALIOP, and evaluate that, in addition to PM2.5? This would remove the need for an assumption of the ratio (taken as 0.6 here), and line 326 notes that there are 409 EPA stations providing both data on a daily basis. Given that this ratio seems to be one of the more uncertain parts of the error budget, it might be that there is more skill in predicting PM10 from CALIOP. Or it might go the other way. That would also be a worthwhile result, since right now we don't know.

Response: We did not estimate PM10 from CALIOP because coarse mode aerosols exhibit vastly different mass extinction efficiency values than those of fine mode aerosols. We have included an initial look into an analysis of coarse mode aerosols, like dust and sea salt, in Table 4, the results of which suggest that large uncertainties would arise for CALIOP-derived PM values assuming coarse mode aerosols as fine mode aerosols. In order to tackle this subject, a more thorough investigation into CALIOP/ground-based aerosol typing is necessary, and we believe this topic is outside the general scope of this paper.

Comment: P17L382: No particular comment here other than to say I am glad that the authors included this specific analysis. It's a point well-made that CALIOP uncertainties propagate downwards so, while CALIOP can see through thin clouds, that does not mean that the data quality is the same as for cloud-free columns.

Response: Thank you for your thoughts on this topic.

Comment: P19L424: This isn't really an uncertainty analysis, so I suggest promoting it from a section 3.2.9 to a section 3.3 by itself. I also have a few suggestions for expansion of this section. It's good to know the correlation lengths across the western vs. eastern USA, but there's a lot of scatter in the plots. Some of this is probably due to limited sampling but some is probably also due to real changes in correlation length. So I wonder if the authors can pull out data from one or two large cities, and one or two remote areas, and highlight the correlation lengths for these (as well as the more general case of east vs. west). This would provide a bit more context about typical correlation lengths in these conditions, which would be helpful for future research built around this analysis.

Response: We agree that Section 3.2.9 is not an uncertainty analysis, and it has been changed to Section 3.3. Concerning the other suggestions, each data point on the plot represents the distance of the given two locations as well as the corresponding PM correlation computed using observations from the two locations. Thus, the datasets are rather discrete and not continuous, as the correlations can only be computed with any two locations with PM observations. Thus, correlation lengths may not be derived reliably using only one or two cities. Still, we emphasize

here that this section is not the focus of the study, and can be explored in a more careful manner in a later paper.

Papers cited:

Chung, A., Chang, D. P., Kleeman, M. J., Perry, K. D., Cahill, T. A., Dutcher, D., ... & Stroud, K: Comparison of real-time instruments used to monitor airborne particulate matter, Journal of the Air & Waste Management Association, 51(1), 109-120, 2001.

Deming, W.E.: Statistical Adjustment of Data, Wiley: New York, 1943.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., and Veihelmann, B.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, J. Geophys. Res.–Atmos., 111, https://doi.org/10.1029/2005JD006619, 2006.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L.: Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge, Atmospheric Environment, 37(9-10), 1277-1292, 2003.

Huang, X. H., Bian, Q., Ng, W. M., Louie, P. K., and Yu, J. Z: Characterization of PM2.5 major components and source investigation in suburban Hong Kong: a one year monitoring study, Aerosol Air Qual. Res, 14(1), 237-250, 2014.

Kiss, G., Imre, K., Molnár, Á., and Gelencsér, A.: Bias caused by water adsorption in hourly PM measurements, Atmos. Meas. Tech., 10, 2477-2484, https://doi.org/10.5194/amt-10-2477-2017, 2017.

Patashnick, H., Rupprecht, G., Ambs, J. L., and Meyer, M. B.: Development of a reference standard for particulate matter mass in ambient air, Aerosol Science & Technology, 34(1), 42-45, 2001.

Spagnolo, G. S.: Automatic instrument for aerosol samples using the beta-particle attenuation, Journal of aerosol science, 20(1), 19-27, 1989.

Toth, T. D., Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Shi, Y., and Westphal, D. L.: Impact of data quality and surface-to-column representativeness on the $PM_{2.5}$ / satellite AOD relationship for the contiguous United States, Atmos. Chem. Phys., 14, 6049-6062, https://doi.org/10.5194/acp-14-6049-2014, 2014.

Response to Anonymous Referee #2

Comment: The scope of the submitted work is to investigate the potential exploitation of CALIOP extinction profiles in order to derive near-surface concentrations of particles with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$). The assessment of the applied methodology is made through the evaluation of the CALIOP derived PM concentrations against corresponding daily ground-based measurements obtained at numerous EPA stations, over the period 2008-2009, distributed across CONUS, which is the area of interest. A powerful element of using vertically resolved retrievals is that the altitude range can be constrained (i.e., near surface where the PM concentrations are measured from the ground) in contrast to passive sensors which are representative for the whole atmospheric column. To my opinion, the issues addressed by the authors fit well to the scientific objectives of AMT and therefore I recommend the submitted manuscript to be published. Nevertheless, I believe that several points must be modified making the text acceptable for publication. My major and minor comments are listed below.

Response: Thank you for your thoughts and encouraging comments.

Comment: The authors have used only 2-year satellite data thus making the robustness of the obtained outcomes questionable taking into account CALIOP's low sampling frequency and narrow footprint. In order to overcome this drawback, you have to repeat the analysis for the full dataset.

Response: We agree that overcoming this sampling drawback can be achieved through extending the analysis for more than two years. However, this would be computationally expensive and is a non-trivial task. We envisioned this manuscript as a proof-of-concept study, the purpose of which is to provide an initial demonstration of the feasibility of our method. Adding other years to the analysis will be one of the focuses of forthcoming CALIOP/PM2.5 papers.

Comment: According to the applied methodology, all the aerosol extinctions assigned as dust in the CALIOP retrieval algorithm are masked out since focus is given on the small size particles (Lines 198-200). However, which is the treatment for the other aerosol subtypes consisting of coarse particles (i.e., marine, dusty marine)? Moreover, what is happening when the aerosol subtype is clean continental? I would suggest to repeat the aerosol type analysis (Section 3.2.8) but considering only the CALIOP aerosol subtypes which are not associated with large size particles (i.e., dust, marine, marine dust) and are relevant to pollution. Keep in mind that appropriate modifications, depending on aerosol types, may be needed in equations 1, 2 and 3 (i.e., mass scattering and absorption efficiencies, hygroscopic growth factor).

Response: All aerosol subtypes not classified as dust are considered for our method (e.g., marine, dusty marine, clean continental, etc.). We have already excluded dust, and most areas of the CONUS are not dominated by sea salt aerosols. Indeed, a statistical analysis showed that CALIOP 100-1000 m aerosol layers consisting entirely of marine (dusty marine) subtypes represent only ~2% (~1%) of all subtypes. Thus, the impact of including these aerosols should be minor. We do note, however, that one of the areas of focus for future studies of CALIOP-derived estimates is a more thorough investigation into aerosol typing.

Comment: Could you please comment why the quality assurance criteria applied here are different than those suggested by Tacket et al. (2018; https://www.atmos-meas-tech.net/11/4129/2018/)?

Response: The QA criteria applied for this paper are the same as those of our previous CALIOP papers (e.g., Toth et al. 2014; 2016; 2018), and we wanted to be consistent with these studies. The QA scheme employed here was developed from Kittaka et al. (2011) and Campbell et al. (2012), both of which provide detailed justifications for the QA choices made. Toth et al. (2016) provides comparisons of aerosol extinction profiles using our QA scheme and those from the CALIPSO Level 3 aerosol profile product. While some differences were found, these were mostly attributed to differences in averaging, treatment of clouds and fill values, and our vertical regridding from 60 m to 100 m.

Comment: Page 7 – Lines 157-160: The inclusion of different PM measurements techniques (filter-based or averages from hourly samples) how can affect the intercomparison results?

Response: As suggested by Kiss et al. (2017), large uncertainties exist in hourly PM data, while less biases are expected for PM data collected over a longer period of time. Thus, there are likely differences in the two methods for collecting PM data. Still, to fully explore this issue requires a study of its own, and thus we have added the following discussions in the text:

"Note that uncertainties have been reported for hourly PM measurements (Kiss et al., 2017). Examples of some uncertainties in these $PM_{2.5}$ measurements depend upon the instrument/method used: gravimetric (e.g., transport to the lab/human error and volatization of PM during the drying process; Patashnick et al., 2001), TEOM (e.g., errors due to improper inlet tube temperature; Eatough et al., 2003), and beta attenuation monitors (e.g., changes in the sample flow rate due to variations in temperature and moisture; Spagnolo, 1989). Also, it has been found that beta attenuation monitors may be more accurate than TEOM, as TEOM can underestimate $PM_{2.5}$ at low temperatures (e.g., Chung et al., 2001). Still, as suggested by Kiss et al. (2017), PM data collected over a longer period of time are much less likely to be biased. Thus, we expect lower uncertainties from data collected over 24-hours, then daily data generated by averaging hourly observations. Fully quantifying the differences from the two different PM observing methods, however, is the subject for a future study."

Comment: Page 4 – Lines 97-102: How much reliable are the scatterplot metrics when MODIS provides daylight AODs while PM concentrations are daily averages? Have you noticed any variation both in spatial and temporal terms?

Response: In this paper, we have not looked into the spatial/temporal variations of MODIS AOD versus PM2.5. This, however, was the subject of one of our past studies (Toth et al., 2014), for which MODIS AOD was compared to both daily (within 1x1 deg.) and hourly (within 40 km) PM2.5 measurements. While larger correlation coefficients were found for the hourly analysis, they still remained low. The purpose of Fig. 1 in this paper was to simply illustrate the limitation of using column-integrated AOD from passive sensors to estimate PM2.5 concentrations near the surface.

Comment: Page 9 – Line 202: A couple of citations are needed here in order to support this argument.

Response: We have added two citations (Nessler et al., 2005 and Lynch et al., 2016), as requested.

Comment: Page 10 – Lines 236-238: It will be useful to provide a map with the number of days participating for the calculation of the average maps illustrated in Figure 3. Moreover, it is required a geographical distribution providing the average number of profiles considered for the derivation of 1° x 1° grid cells (i.e. an indicator of spatial representativeness within the 1deg grid cell).

Response: Thank you for this suggestion. We have added the requested maps as a figure in an appendix. Also, the following description was added to the text in Section 3.1: "Note that, for context, maps of the number of days and CALIOP Level 2 5 km aerosol profiles used in the creation of Fig. 3a-d are shown in Appendix Fig. 1."

Comment: Page 12 – Lines 270-279: I don't agree with the collocation criteria applied here. The horizontal distance (100 km) between CALIOP and PM station probably is too long since the analysis focuses on $PM_{2.5}$ originating from pollution. Under these cases it is expected that the horizontal variability will be very strong and the concentrations will decrease rapidly for increasing distance from the source. As it concerns the temporal collocation, the optimum solution would be to use PM measurements available at the finest temporal resolution thus making feasible an appropriate matching with the CALIOP near-surface profiles. On the contrary, if the ground-based data are provided only as daily averages then you cannot consider that a satellite overpass and a daily average are temporally collocated. In the former data you have an instantaneous observation while in the latter one the diurnal variation is included. In case where the EPA data are given only on a daily basis, then it is more convenient to compare "daily" CALIOP profiles (considering dates where both the daytime and nighttime satellite retrievals are available) against the corresponding surface PM_{10} concentrations. For this reason, I believe that Figures 3-e and 3-f as well as the relevant parts of the text must be removed. Please consider this comment throughout your analysis.

Response: Thanks for the suggestion. We agree that "daily" averaged CALIOP profiles may be used for comparing with daily averaged surface PM observations. However, with a narrow swath of ~70 m and a repeat cycle of 16 days, very few data points would be available within 100 km of a particular EPA site for both daytime and nighttime CALIOP aerosol profiles. For the spatial collocation, the +/- 100 km collocation distance is used here, as we considered the spread of aerosols within 24 hours. For example, for a 10 km/hour wind speed, aerosol particles may travel 200 km (or +/- 100 km) within 24 hours. Also, as suggested from this paper, the averaged e-folding correlation length for PM2.5 concentrations over the CONUS is ~600 km, and thus we believe 100 km is a reasonable collocation range.

Also, analysis using finer temporal resolution PM2.5 data may produce better results under some conditions, but comes with its own issues. For example, there are insufficient collocated CALIOP profiles and hourly PM2.5 data over a two-year period for the CONUS, so the temporal domain would need to be greatly expanded. Secondly, this type of study would take careful analysis of the CALIOP data, as individual CALIOP aerosol extinction profiles could be subject to higher

uncertainties (e.g., rather than using a two-year mean). These research topics will be examined in detail in future studies.

Comment: Section 3.2.1: Considering my previous comment, the analysis should be presented only for the "daily" CALIOP – PM pairs and not separately for daytime and nighttime. Likewise, the CALIOP derived $PM_{2.5}$ ranges (x axis in Figure 5) should be equally sampled and not grouped based on user-defined bins of PM concentrations. In addition, the authors are stating in Lines 314-316 that the computations have not been done for PM concentrations ≥ 25 μg m^{-3} due to the limited number of concurrent annual means. However, according to Figure 5, the number of samples for the lowest bin (< 5 μg m^{-3}) during daytime is almost zero (the same is valid for the highest bins, particularly for the nighttime retrievals). Is that correct? Can we trust the calculated RMSEs resulting from a very small number of samples?

Response: As mentioned in another response, using only "daily" CALIOP-PM pairs is not feasible here for a robust analysis, due to the repeat cycle of the CALIPSO satellite. Very few data points would be available within 100 km of a particular EPA site for both daytime and nighttime CALIOP aerosol profiles. Thus, we leave daytime and nighttime separated for this figure. As for the CALIOP derived PM2.5 ranges, we have adjusted them such that each bin is equally sampled based upon a cumulative histogram analysis. Each point from left to right in the new Fig. 5 represents the RMSE and mean PM2.5 concentration derived from CALIOP for 0-20%, 20-40%, 40-60%, 60-80%, and 80-100% cumulative frequencies. This addresses the other items in this comment, like those of few samples for the lowest and highest bins in the old Fig. 5. Because they are now equally sampled, we have removed the secondary y-axis since the number of samples do not change as a function of CALIOP-derived PM2.5 concentration. We have also revised the corresponding text in Section 3.2.1 as follows: "As a first step for the uncertainty analysis, we estimated the prognostic error of 2-year averaged $PM_{2.5\_CALIOP}$. Figure 5 shows the root-mean-square error (RMSE) of CALIOP-based $PM_{2.5}$ concentrations against those from EPA stations as a function of CALIOP-based $PM_{2.5}$ for the 2008-2009 period over the CONUS. RMSEs were computed for five equally sampled bins, determined from a cumulative histogram analysis. Each point in Fig. 5, from left to right, represents the RMSE and mean $PM_{2.5}$ concentration derived from CALIOP for 0-20%, 20-40%, 40-60%, 60-80%, and 80-100% cumulative frequencies. A mean combined daytime and nighttime RMSE of ~4 μg m^{-3} is found, with a mean value slightly greater for nighttime (~4.3 μg m^{-3}) than daytime (~3.7 μg m^{-3}). While most bins exhibit larger nighttime RMSEs, daytime RMSEs are larger for the greatest mean CALIOP-derived $PM_{2.5}$ concentrations."

Comment: Section 3.2.2: To my opinion this sensitivity study should be the first step of the analysis in order to define the most "representative" altitude range. According to the summary statistics presented in Table 2, it seems that it is better to restrict the upper bound at 600 – 700m.

Response: While a surface layer up to about 600-700 m results in larger r^2 values, much variability in the statistics exists between surface layer heights (as shown in Table 2). Also, differences are found between daytime and nighttime for various layers. One possible issue is a lower signal-to-noise ratio if we restricted the surface layer to lower heights. We stress that the purpose of this paper is an initial exploration of the topic, and wanted to include Table 2 as a first look at surface layer height sensitivity. Another study is necessary to better evaluate this subject, especially as surface layer height changes regionally and diurnally.

Comment: Section 3.2.4: Which is the impact on the r^2 values?

Response: The r^2 values are not impacted by varying the PM2.5/PM10 ratio. This is because all of the CALIOP-derived PM2.5 points for each scenario shown in Table 3 are multiplied by a common ratio (see Equation 3), but the collocated EPA concentrations remain unchanged (thus not altering the correlation).

Comment: Section 3.2.5: Instead of presenting daytime and nighttime CALIOP derived PM concentrations it is better to consider only the daily (computed from the concurrent daytime and nighttime profiles) ones (see comment 6).

Response: Thank you for the suggestion. We believe it is important to show the daytime and nighttime analyses separately, and an analysis using concurrent daytime and nighttime profiles collocated with a particular EPA site will not yield many samples due to the repeat cycle of the CALIPSO satellite. Thus we didn't make the change.

Comment: Page 19 – Lines 448-450: This means that the CALIOP derived PM concentrations are not reliable in coastal (contamination by sea-salt particles) or dust affected regions?

Response: The large uncertainties are because mass extinction efficiencies are drastically different for coarse and fine mode aerosols. Here we applied mass extinction efficiencies from fine mode aerosols to coarse mode aerosols, and not surprisingly, see large uncertainties. Lower uncertainties can be expected if we apply coarse mode mass extinction efficiencies to coarse mode aerosols. However, this puts the pressure on accurate estimations of aerosol types from CALIOP or other lidar observations, which we believe is a study of its own, and will be investigated in future studies.

Comment: Section 3.2.9: In this section it would be also useful to provide a map with the distances where the 1/e value is found at each station.

Response: Thank you for this suggestion. However, for each pair of PM observing locations, one correlation value is computed for a given distance between the two locations. Thus, the analysis is discrete, not continuous. The 1/e values are estimated from Fig. 10, which is composed of individual points representing both a distance and spatial PM2.5 correlation between pairs of EPA sites over the CONUS. If we apply the same analysis to a given PM observing location, it is likely to have data gaps due to the discrete nature of the dataset. Thus, we leave Fig. 10 untouched.

Comment: Page 3 – Lines 81-84: Could you please explain better this sentence?

Response: We have rewritten the sentence to:

"Indeed, Kaku et al. (2018) recently showed that surface PM2.5 had longer spatial correlation lengths than AOD, even in the "well behaved" southeastern United States where previous studies showed good correlation between PM2.5 and AOD (e.g., Wang and Christopher, 2003)."

Comment: Page 4 – Lines 91-94: It is not clear what the authors want to say here.

Response: We have rewritten the sentence to:

"It is arguable that from a climatological/long-term average perspective, the use of AOD as a proxy for PM2.5 concentrations nevertheless has certain qualitative skill (e.g., Toth et al., 2014; Reid et al., 2017) due to the averaging process that suppresses sporadic aerosol events with highly variable vertical distributions."

Comment: Page 10 – Line 244: What do you mean exactly here? ("..., as surface layer heights may change seasonally and diurnally.")

Response: We have removed "as surface layer heights may change seasonally and diurnally" to avoid confusion.

Comment: Page 19 – Line 431: Sulfate & organic or just sulfate?

Response: To avoid confusion, we removed "& organic". But primary and secondary biogenic aerosols are mostly fine mode as well.

Comment: Page 20 – Lines 456-458: Please rephrase this sentence.

Response: This sentence was broken into two sentences, as follows: "To accomplish this, all EPA stations over the CONUS with at least 50 days of daily data available for the 2008-2009 period were first determined. Next, the distances between each pair of these EPA stations, and their corresponding correlation of daily $PM_{2.5}$ concentrations, were computed."

Papers cited:

Campbell, J. R., Tackett, J. L., Reid, J. S., Zhang, J., Curtis, C. A., Hyer, E. J., Sessions, W. R., Westphal, D. L., Prospero, J. M., Welton, E. J., Omar, A. H., Vaughan, M. A., and Winker, D. M.: Evaluating nighttime CALIOP 0.532 μm aerosol optical depth and extinction coefficient retrievals, Atmos. Meas. Tech., 5, 2143-2160, https://doi.org/10.5194/amt-5-2143-2012, 2012.

Chung, A., Chang, D. P., Kleeman, M. J., Perry, K. D., Cahill, T. A., Dutcher, D., ... & Stroud, K: Comparison of real-time instruments used to monitor airborne particulate matter, Journal of the Air & Waste Management Association, 51(1), 109-120, 2001.

Deming, W.E.: Statistical Adjustment of Data, Wiley: New York, 1943.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L.: Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge, Atmospheric Environment, 37(9-10), 1277-1292, 2003.

Huang, X. H., Bian, Q., Ng, W. M., Louie, P. K., and Yu, J. Z: Characterization of PM2.5 major components and source investigation in suburban Hong Kong: a one year monitoring study, Aerosol Air Qual. Res, 14(1), 237-250, 2014.

Kiss, G., Imre, K., Molnár, Á., and Gelencsér, A.: Bias caused by water adsorption in hourly PM measurements, Atmos. Meas. Tech., 10, 2477-2484, https://doi.org/10.5194/amt-10-2477-2017, 2017.

Kittaka, C., Winker, D. M., Vaughan, M. A., Omar, A., & Remer, L. A.: Intercomparison of column aerosol optical depths from CALIPSO and MODIS-Aqua, Atmospheric Measurement Techniques, 4(2), 131, https://doi.org/10.5194/amt-4-131-2011, 2011.

Lynch P., and coauthors: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences, Geosci. Model Dev., 9, 1489-1522, doi:10.5194/gmd-9-1489-2016, 2016.

Nessler, R., Weingartner, E., and Baltensperger, U. (2005). Effect of humidity on aerosol light absorption and its implications for extinction and the single scattering albedo illustrated for a site in the lower free troposphere, *Journal of Aerosol Science*, *36*(8), 958-972.

Patashnick, H., Rupprecht, G., Ambs, J. L., and Meyer, M. B.: Development of a reference standard for particulate matter mass in ambient air, Aerosol Science & Technology, 34(1), 42-45, 2001.

Spagnolo, G. S.: Automatic instrument for aerosol samples using the beta-particle attenuation, Journal of aerosol science, 20(1), 19-27, 1989.

Toth, T. D., Campbell, J. R., Reid, J. S., Tackett, J. L., Vaughan, M. A., Zhang, J., and Marquis, J. W.: Minimum aerosol layer detection sensitivities and their subsequent impacts on aerosol optical thickness retrievals in CALIPSO level 2 data products, Atmos. Meas. Tech., 11, 499-514, https://doi.org/10.5194/amt-11-499-2018, 2018.

Toth, T. D., Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Shi, Y., and Westphal, D. L.: Impact of data quality and surface-to-column representativeness on the $PM_{2.5}$ / satellite AOD relationship for the contiguous United States, Atmos. Chem. Phys., 14, 6049-6062, https://doi.org/10.5194/acp-14-6049-2014, 2014.

Toth, T. D., Zhang, J., Campbell, J. R., Reid, J. S., and Vaughan, M. A.: Temporal variability of aerosol optical thickness vertical distribution observed from CALIOP, Journal of Geophysical Research: Atmospheres, 121(15), 9117-9139, https://doi.org/10.1002/2015JD024668, 2016.

A bulk-mass-modeling-based method for retrieving Particulate Matter Pollution using CALIOP observations

Travis D. Toth[1], Jianglong Zhang[2], Jeffrey S. Reid[3], and Mark A. Vaughan[1]

[1]NASA Langley Research Center, Hampton, VA
[2]Department of Atmospheric Sciences, University of North Dakota, Grand Forks, ND
[3]Marine Meteorology Division, Naval Research Laboratory, Monterey, CA

Correspondence to: Travis D. Toth (travis.d.toth@nasa.gov); Jianglong Zhang (jianglong.zhang@und.edu)

Abstract. In this proof-of-concept paper, we apply a bulk-mass-modeling method using observations from the NASA Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument for retrieving particulate matter (PM) concentration over the contiguous United States (CONUS) over a 2-year period (2008-2009). Different from previous approaches that rely on empirical relationships between aerosol optical depth (AOD) and $PM_{2.5}$ (PM with particle diameters less than 2.5 μm), for the first time, we derive $PM_{2.5}$ concentrations, both at daytime and nighttime, from near surface CALIOP aerosol extinction retrievals using bulk mass extinction coefficients and model-based hygroscopicity. Preliminary results from this 2-year study conducted over the CONUS show a good agreement (r^2 ~ 0.48; mean bias of -3.3 μg m^{-3}) between the averaged nighttime CALIOP-derived $PM_{2.5}$ and ground-based $PM_{2.5}$ (with a lower r^2 of ~0.21 for daytime; mean bias of -0.4 μg m^{-3}), suggesting that PM concentrations can be obtained from active-based spaceborne observations with reasonable accuracy. Results from sensitivity studies suggest that accurate aerosol typing is needed for applying CALIOP measurements for $PM_{2.5}$ studies. Lastly, the e-folding correlation length for surface $PM_{2.5}$ is found to be around 600 km for the entire CONUS (~300 km for Western CONUS and ~700 km for Eastern CONUS), indicating

Deleted: sizes

that CALIOP observations, although sparse in spatial coverage, may still be applicable for $PM_{2.5}$ studies.

1 Introduction

During the last decade, an extensive number of studies have researched the feasibility of estimating $PM_{2.5}$ (particulate matter with particle diameters smaller than 2.5 μm) pollution with the use of passive-based satellite-derived aerosol optical depth (AOD; e.g., Liu et al., 2007; Hoff and Christopher, 2009; van Donkelaar et al., 2015). Monitoring of PM concentration from space observations is needed, as $PM_{2.5}$ pollution is one of the known causes of respiratory related diseases as well as other health related issues (e.g., Liu et al., 2005; Hoff and Christopher, 2009; Silva et al., 2013). Yet, ground-based $PM_{2.5}$ measurements are often inconsistent or have limited availability over much of the globe.

In some earlier studies, empirical relationships of $PM_{2.5}$ concentrations and AODs were developed and used for estimating $PM_{2.5}$ concentrations from passive sensor retrieved AODs (e.g., Wang and Christopher, 2003; Engel-Cox et al., 2004; Liu et al., 2005; Kumar et al., 2007; Hoff and Christopher, 2009). One of the limitations of this approach is that vertical distributions and thermodynamic state of aerosol particles vary with space and time. Especially for regions with elevated aerosol plumes, deep boundary layer entrainment zones, or strong nighttime inversions, column-integrated AODs are not a good approximation of surface $PM_{2.5}$ concentrations at specific points and times (e.g., Liu et al., 2004; Toth et al., 2014; Reid et al., 2017). Indeed, Kaku et al. (2018) recently showed that surface $PM_{2.5}$ had longer spatial correlation lengths than AOD, even in the "well behaved" southeastern United States where previous studies showed good correlation between $PM_{2.5}$ and AOD (e.g., Wang and Christopher, 2003). To account for variability in aerosol

Deleted: sizes

Deleted: performance

vertical distribution, several studies have attempted the use of chemical transport models, or CTMs (e.g., van Donkelaar et al., 2015). Satellite data assimilation of AOD has become commonplace, vastly improving AOD analyses and short-term prediction (e.g., Zhang et al., 2014; Sessions et al., 2015). Yet, $PM_{2.5}$ simulations remain poor (e.g., Reid et al., 2016). Uncertainties in such studies are unavoidable due to uncertainties in CTM-based aerosol vertical distributions, and no nighttime AODs are currently available from passive-based satellite retrievals.

It is arguable that from a climatological/long-term average perspective, the use of AOD as a proxy for $PM_{2.5}$ concentrations nevertheless has certain qualitative skill (e.g., Toth et al., 2014; Reid et al., 2017) due to the averaging process that suppresses sporadic aerosol events with highly variable vertical distributions. Still, as illustrated in Fig. 1, where 2-year (2008-2009) means of Moderate Resolution Imaging Spectroradiometer (MODIS) AOD are plotted against $PM_{2.5}$ concentrations throughout the contiguous United States (CONUS), although a linear relationship is plausibly shown, a low r^2 value of 0.08 is found. To construct Fig. 1, Aqua MODIS Collection 6 (C6) Optical_Depth_Land_And_Ocean data (0.55 μm), restricted to "Very Good" retrievals as reported by the Land_Ocean_Quality_Flag, are first collocated with daily surface $PM_{2.5}$ measurements in both space and time (i.e., within 40 km in distance and the same day), and then collocated daily pairs are averaged into 2-year means (for each $PM_{2.5}$ site). Figure 1 may be indicating that even from a long-term mean perspective, aerosol vertical distributions are not uniform across the CONUS, which is also confirmed by other studies (e.g., Toth et al., 2014). AOD retrievals themselves, with known uncertainties due to cloud contamination and assumptions in the retrieval process (e.g., Levy et al., 2013), may also introduce uncertainties to that task.

On board the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite, the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument

Deleted: for the most significant events as well as

Deleted: 9

provides observations of aerosol and cloud vertical distributions at both day and night (Hunt et al., 2009; Winker et al., 2010). Given that CALIOP provides aerosol extinction retrievals near the ground, it is interesting and reasonable to raise the question: can near surface CALIPSO extinction be used as a better physical quantity than AOD for estimating surface $PM_{2.5}$ concentrations? This is because unlike AOD, which is a column-integrated value, near surface CALIPSO extinction is, in theory, a more realistic representation of near surface aerosol properties. Yet, in comparing with passive sensors such as MODIS, which has a swath width on the order of ~2000 km, CALIOP is a nadir pointing instrument with a narrow swath of ~70 m and a repeat cycle of 16 days (Winker et al., 2009). Thus, the spatial sampling of CALIOP is sparse on a daily basis and temporal sampling or other conditional or contextual biases are unavoidable if CALIOP observations are used to estimate daily $PM_{2.5}$ concentrations (Zhang and Reid, 2009; Colarco et al., 2014). Also, there are known uncertainties in CALIPSO retrieved extinction values due to uncertainties in the retrieval process, such as the lidar ratio (extinction-to-backscatter ratio), calibration, and the "retrieval fill value" (RFV) issue (Young et al., 2013; Toth et al., 2018).

Even with these known issues, especially the sampling bias, it is still compelling to investigate if near surface CALIOP extinction can be utilized to retrieve surface $PM_{2.5}$ concentrations with reasonable accuracy from a long-term (i.e., two-year) mean perspective. CALIOP data have been successfully used in $PM_{2.5}$ studies in the past, but primarily for assisting passive-based AOD/$PM_{2.5}$ analyses using aerosol vertical distribution as a constraint (e.g., Glantz et al., 2009; van Donkelaar et al., 2010; Val Martin et al., 2013; Toth et al., 2014; Li et al., 2015; Gong et al., 2017). However, the question remained as to the efficacy of the direct use of CALIOP retrievals. To demonstrate a concept, we developed a bulk mass scattering scheme for inferring PM concentrations from near surface aerosol extinction retrievals derived from CALIOP

observations. The bulk method used here is based upon the well-established relationship between particle light scattering and $PM_{2.5}$ aerosol mass concentration (e.g., Charlson et al., 1968; Waggoner and Weiss, 1980; Liou, 2002; Chow et al., 2006), discussed further, with the relevant equations, in Sect. 2.

In this study, using two years (2008-2009) of CALIOP and United States (U.S.) Environmental Protection Agency (EPA) data over the CONUS, the following questions are addressed:

1. Can CALIOP extinction be used effectively for estimating $PM_{2.5}$ concentrations through a bulk mass scattering scheme from a 2-year mean perspective for both daytime and nighttime?
2. Can CALIOP extinction be used as a better parameter than AOD for estimating $PM_{2.5}$ concentrations from a 2-year mean perspective?
3. What are the sampling biases we can expect in CALIOP estimates of $PM_{2.5}$?
4. How do uncertainties in bulk properties compare to overall CALIOP-retrieved $PM_{2.5}$ uncertainty?

Details of the methods and datasets used are described in Sect. 2. Section 3 shows the preliminary results using two years of EPA $PM_{2.5}$ and CALIOP data, including an uncertainty analysis. The conclusions of this paper are provided in Sect. 4.

2 Data and Methods

Since 1970, the U.S. EPA has monitored surface PM using a number of Federal Reference/Equivalent Methods (FRMs/FEMs), which employ gravimetric, tapered element oscillating microbalance (TEOM), and beta gauge instruments (Federal Register, 1997; Greenstone, 2002). Two years (2008-2009) of daily $PM_{2.5}$ Local Conditions (EPA code = 88101) data were acquired from the EPA Air Quality System for use in this investigation, consistent with our previous $PM_{2.5}$ study (Toth et al., 2014). These data represent $PM_{2.5}$ concentrations over a 24-hour period and include two scenarios: one sample is taken during the 24-hour duration (i.e., filter-based measurement), or an average is computed from hourly samples within this time period (every hour may not have an available measurement, however).

Note that uncertainties have been reported for hourly PM measurements (Kiss et al., 2017). Examples of some uncertainties in these $PM_{2.5}$ measurements depend upon the instrument/method used: gravimetric (e.g., transport to the lab/human error and volatization of PM during the drying process; Patashnick et al., 2001), TEOM (e.g., errors due to improper inlet tube temperature; Eatough et al., 2003), and beta attenuation monitors (e.g., changes in the sample flow rate due to variations in temperature and moisture; Spagnolo, 1989). Also, it has been found that beta attenuation monitors may be more accurate than TEOM, as TEOM can underestimate $PM_{2.5}$ at low temperatures (e.g., Chung et al., 2001). Still, as suggested by Kiss et al. (2017), PM data collected over a longer period of time are much less likely to be biased. Thus, we expect lower uncertainties from data collected over 24-hours, then daily data generated by averaging hourly observations. Fully quantifying the differences from the two different PM observing methods, however, is the subject for a future study.

CALIOP, flying aboard the CALIPSO platform within the A-Train satellite constellation, is a dual wavelength (0.532 and 1.064 μm) lidar that has collected profiles of atmospheric aerosol

Deleted: We note that

Deleted: t

Deleted: Note that uncertainties have been reported for hourly PM measurements (Kiss et al., 2017). Examples of some uncertainties in these $PM_{2.5}$ measurements depend upon the instrument/method used: gravimetric (e.g., transport to the lab/human error and volatization of PM during the drying process; Patashnick et al., 2001), TEOM (e.g., errors due to improper inlet tube temperature; Eatough et al., 2003), and beta gauge (e.g., changes in the sample flow rate due to variations in temperature and moisture; Spagnolo, 1989). Also, it has been found that BAM may be more accurate than TEOM, as TEOM can underestimate $PM_{2.5}$ at low temperatures (e.g., Chung et al., 2001). Still, as suggested by Kiss et al. (2017), PM data collected over a longer period of time are much less likely to be biased. Thus, we expected less uncertainties from data collected over 24-hours, then daily data generated by averaging hourly observations. Fully quantifying the differences from the two different PM observing methods, however is subject to a future study. ¶

particles and clouds since summer 2006 (Winker et al., 2007). In this study, daytime and nighttime extinction coefficients retrieved at 0.532 µm from the Version 4.10 CALIOP Level 2 5 km aerosol profile (L2_05kmAPro) product were used. Using parameters provided in the L2_05kmAPro product, as well as the corresponding Level 2 5 km aerosol layer (L2_05kmALay) product, a robust quality-assurance (QA) procedure for the aerosol observations was implemented (Table 1). Further information on the QA metrics and screening protocol are discussed in detail in previous studies (Kittaka et al. 2011; Campbell et al. 2012; Toth et al. 2013; 2016). Once the QA procedure was applied, the aerosol profiles were linearly re-gridded from 60 m vertical resolution (above mean sea level [AMSL]) to 100 m segments (i.e., resampled to 100 m resolution) referenced to the local surface (above ground level [AGL]; Toth et al., 2014; 2016). The choice of 100 m was arbitrary, and the profiles were re-gridded in order to obtain an AGL-corrected dataset, as opposed to the AMSL-referenced profiles provided by the L2_05kmAPro product. Surface elevation and relative humidity (RH) were taken from collocated model data included in the CALIPSO L2_05kmAPro product (CALIPSO Data Products Catalog (Release 4.20); RH taken from the Modern Era Retrospective-Analysis for Research, or MERRA-2 reanalysis product). To limit the effects of signal attenuation and increase the chances of measuring aerosol presence near the surface, the Atmospheric Volume Description parameter within the L2_05kmAPro dataset is used to cloud-screen each aerosol profile as in Toth et al. (2018).

In this study, near surface PM mass concentration (C_m) is derived from near surface CALIOP extinction based on a bulk formulation as in Equation 1 (e.g., Liou, 2002; Chow et al., 2006):

$$\beta = C_m(a_{scat}f_{rh} + a_{abs}) \text{ x } 1000 \qquad (1)$$

where β is CALIOP-derived near surface extinction in km^{-1}, C_m is the PM mass concentration in µg m^{-3}, a_{scat} and a_{abs} are dry mass scattering and absorption efficiencies in m^2 g^{-1}, and f_{rh} represents the light scattering hygroscopicity, respectively. As a preliminary study, for the purpose of demonstrating this concept, we assume the dominant aerosol type over the contiguous U.S. (CONUS) is pollution aerosol (i.e., the most prevalent near-surface aerosol type reported in the CALIOP products for the CONUS during 2008-2009 is polluted continental) with a_{scat} and a_{abs} values of 3.40 and 0.37 m^2 g^{-1} (Hess et al., 1998; Lynch et al., 2016), respectively. These values are similar to those reported in Malm and Hand (2007) and Kaku et al. (2018) but are interpolated to 0.532 µm from values at 0.450 µm and 0.550 µm obtained from the Optical Properties of Aerosols and Clouds (OPAC) model (Hess et al., 1998). Still, both a_{scat} and a_{abs} have regional and species related dependencies. Also, only 2-year averages are used in this study, and we assume that sporadic aerosol plumes are smoothed out in the averaging process, and that bulk aerosol properties are similar throughout the study region. We have further explored the impact of aerosol types to $PM_{2.5}$ retrievals in a later section. Furthermore, to aid in focusing this study on fine mode/anthropogenic aerosols, those aerosol extinction range bins classified as dust by the CALIOP typing algorithm were excluded from the analysis.

Also, surface PM concentrations are dry mass measurements. To account for the impact of humidity on a_{scat} (it is assumed that a_{abs} is not affected by moisture; Nessler et al., 2005; Lynch et al., 2016), we estimated the hygroscopic growth factor for pollution aerosol based on Hanel (1976), as shown in Equation 2:

$$f_{rh} = \left(\frac{1-RH}{1-RH_{ref}}\right)^{-\Gamma} \qquad (2)$$

where f_{rh} is the hygroscopic growth factor, RH is the relative humidity, and RH_{ref} is the reference RH and is set to 30% in this study (Lynch et al., 2016). Γ is a unitless value (a fit parameter describing the amount of hygroscopic increase in scattering) and is assumed to be 0.63 (i.e., sulfate aerosol) in this study (Hanel, 1976; Chew et al., 2016; Lynch et al., 2016).

Additionally, the CALIOP-derived PM density is for all particle sizes. To convert from mass concentration of PM (C_m) to mass concentration of $PM_{2.5}$ ($C_{m2.5}$), which represents mass concentration for particle diameters smaller than 2.5 µm, we adopted the $PM_{2.5}$ to PM_{10} (PM with diameters less than 10 µm) ratio (ϕ) of 0.6 as measured during the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC^{4}RS) campaign over the US (Kaku et al., 2018). Again, the ratio of $PM_{2.5}$ to PM_{10} can also vary spatially, however we used a regional mean to demonstrate the concept. Analyses in a later section using two-years (2008-2009) of surface $PM_{2.5}$ to PM_{10} data suggest that 0.6 is a rather reasonable number to use for the CONUS for the study period. Here we assume that mass concentrations for particle diameters larger than 10 µm are negligible over the CONUS. Thus, we can rewrite Equation 1 as:

$$C_{m2.5} = \frac{\beta \text{ x } \phi}{(a_{scat} \text{ x } f_{rh} + a_{abs}) \text{ x } 1000} \qquad (3)$$

where $C_{m2.5}$ is the CALIOP-derived $PM_{2.5}$ concentration in units of µg m^{-3}.

Lastly, we note that most of the results are shown in the form of scatter plots with fits from Deming regression (Deming, 1943). Due to uncertainties in $PM_{2.5}$ data, we show slopes computed from Deming regression analyses instead of those from simple linear regression. Deming regression in particular is appropriate here, as it accounts for errors in both the independent and dependent variables (Deming, 1943), and has been used in past $PM_{2.5}$ related studies (e.g., Huang et al., 2014).

Deleted: Lastly

Deleted: sizes

Deleted: sizes

3 Results and Discussion

3.1 Regional analysis

Figure 2a shows the mean $PM_{2.5}$ concentration using two years (2008-2009) of daily surface $PM_{2.5}$ data from the U.S. EPA ($PM_{2.5\_EPA}$), not collocated with CALIOP observations. A total of 1,091 stations (some operational throughout the entire period; others only partially) are included in the analysis and observations from those stations are further used in evaluating CALOP-derived $PM_{2.5}$ concentrations ($C_{m2.5}$), as later shown in Fig. 3. $PM_{2.5}$ concentrations of ~10 μg m^{-3} are found over the eastern CONUS. In comparison, much lower $PM_{2.5}$ concentrations of ~5 μg m^{-3} are exhibited for the interior CONUS, over states including Montana, Wyoming, North Dakota, South Dakota, Utah, Colorado, and Arizona. For the west coast of the CONUS, and especially over California, higher $PM_{2.5}$ concentrations are observed, with the maximum two-year mean near 20 μg m^{-3}. Note that the spatial distribution of surface $PM_{2.5}$ concentrations over the CONUS as shown in Fig. 2a is consistent with reported values from several studies (e.g., Hand et al., 2013; Van Donkelaar et al., 2015; Di et al., 2017).

Figure 3a shows the two-year averaged 1° x 1° (latitude/longitude) gridded daytime CALIOP aerosol extinction over the CONUS using CALIOP observations from 100-1000 m, referenced to the number of cloud-free L2_05kmAPro profiles in each 1 x 1° bin. The lowest 100 m of CALIOP extinction data are not used in the analysis due to the potential of surface return contamination (e.g., Toth et al., 2014), although this has been improved for the Version 4 CALIOP products but may still be present in some cases. Here the averaged extinction from 100-1000 m is used to represent near surface aerosol extinction. This selection of the 100-1000 m layer is somewhat arbitrary, even though it is estimated from the mean CALIOP-based aerosol vertical distribution over the CONUS (Toth et al., 2014). Thus, a sensitivity study is provided in a later

Deleted: , as surface layer heights may change seasonally and diurnally

section to understand the impact of this aerosol layer selection to CALIOP-based $PM_{2.5}$ retrievals. As shown in Fig. 3a, higher mean near surface CALIOP extinction of 0.1 km^{-1} are found for the eastern CONUS and over California, while lower values of 0.025-0.05 km^{-1} found for the interior CONUS. Figure 3b shows a plot similar to Fig. 3a but using nighttime CALIOP observations only. Although similar spatial patterns are found during both day and night, the near surface extinction values are overall lower for nighttime than daytime, and nighttime data are less noisy than daytime. These findings are not surprising, as daytime CALIOP measurements are subject to contamination from background solar radiation (e.g., Omar et al., 2013).

To investigate any diurnal biases in the data, Figs. 3c and 3d show the derived $PM_{2.5}$ concentration using daytime and nighttime CALIOP data respectively, based on the method described in Section 2. Both Figures 3c and 3d suggest a higher $PM_{2.5}$ concentration of ~10-12.5 μg m^{-3} over the eastern CONUS, and a much lower $PM_{2.5}$ concentration of ~2.5-5 μg m^{-3} over the interior CONUS. High $PM_{2.5}$ values of 10-20 μg m^{-3} are also found over the west coast of the CONUS, particularly over California. The spatial distribution of $PM_{2.5}$ concentrations, as derived using near surface CALIOP data (Figs. 3c and 3d, as well as the combined daytime and nighttime perspective shown in Fig. 2c), is remarkably similar to the spatial distribution of $PM_{2.5}$ values as estimated based on ground-based observations (Fig. 2a). Still, day and night differences in $PM_{2.5}$ concentrations are also clearly visible, as higher $PM_{2.5}$ values are found, in general, during daytime, based on CALIOP observations. The high daytime $PM_{2.5}$ values, as shown in Fig. 3c, may represent stronger near surface convection and more frequent anthropogenic activities during daytime. However, they may also be partially contributed from solar radiation contamination. Another possibility is that the daytime mean extinction coefficients (from which the mean $PM_{2.5}$ estimates are derived) appear artifically larger than at night due to high daytime noise limiting the

ability of CALIOP to detect fainter aerosol layers during daylight operations. Note that, for context, maps of the number of days and CALIOP Level 2 5 km aerosol profiles used in the creation of Fig. 3a-d are shown in Appendix Fig. 1.

Figure 3e shows the inter-comparison between $PM_{2.5\_EPA}$ and $PM_{2.5\_CALIOP}$ concentrations. Note that only CALIOP and ground-based $PM_{2.5}$ data pairs, which are within 100 km of each other and have reported values for the same day (i.e., year, month, and day), are used to generate Fig. 3e. Still, although only spatially and temporally collocated data pairs are used, ground-based $PM_{2.5}$ data represent 24-hour averages, while CALIOP-derived $PM_{2.5}$ concentrations are instantaneous values over the daytime CALIOP overpass. To reduce this temporal bias, two years (2008-2009) of collocated CALIOP-derived and measured $PM_{2.5}$ concentrations are averaged and only the two-year averages are used in constructing Fig 3e. Also, to minimize the above-mentioned temporal sampling bias, ground stations with fewer than 100 collocated pairs are discarded. This leaves a total of 276 stations for constructing Fig. 3e.

As shown in Fig. 3e, an r^2 value of 0.21 (with slope of 1.07) is found between CALIOP-derived and measured surface $PM_{2.5}$ concentrations, with a corresponding mean bias of -0.40 μg m^{-3} ($PM_{2.5\_CALIOP}$ - $PM_{2.5\_EPA}$). In comparison, Fig. 3f shows results similar to Fig. 3e, but for nighttime CALIOP data. A much higher r^2 value of 0.48 (with slope of 0.96) is found between CALIOP-derived and measurement $PM_{2.5}$ values from 528 EPA stations, with a corresponding mean bias of -3.3 μgm^{-3} ($PM_{2.5\_CALIOP}$ - $PM_{2.5\_EPA}$). This may be related to the diurnal variability of $PM_{2.5}$ concentrations, as the daily mean EPA measurement might be closer to the CALIOP A.M. retrieval than to its P.M. counterpart. Still, data points are more scattered in Fig. 3e in comparison with Fig. 3f, which again indicates that daytime CALIOP data are noisier, possibly due to daytime

Deleted: a

Deleted: 0.48

Deleted: a

Deleted: 0.67

solar contamination as well as other factors such as biases in relative humidity. Details of these biases are further explored in Section 3.2.

To supplement this analysis, a pairwise $PM_{2.5\_EPA}$ and $PM_{2.5\_CALIOP}$ (day and night CALIOP combined) analysis is presented in the spatial plots of Figs. 2b and 2d. Here, however, we lift the 100 collocated pairs requirement to increase data samples for better spatial representativeness. The spatial variability of $PM_{2.5}$ over the CONUS is consistent with the observed patterns of non-collocated data (i.e., Figs. 2a and 2c), but with generally higher values due to differences in sampling. Also, comparing Figs. 2b and 2d, $PM_{2.5\_EPA}$ spatial patterns match well with those of $PM_{2.5\_CALIOP}$, yet with larger values for $PM_{2.5\_EPA}$ (consistent with the biases discussed above). Lastly, a scatterplot of the pairwise analysis shown in Figs. 2b and 2d is provided in Fig. 4. An r^2 value of 0.40 is found between EPA and CALIOP-derived $PM_{2.5}$ concentrations from a combined daytime and nighttime CALIOP perspective. Overall, Figs. 2, 3, and 4 indicate that near surface CALIOP extinction data can be used to estimate surface $PM_{2.5}$ concentrations with reasonable accuracy.

3.2 Uncertainty analysis

In this section, uncertainties in the CALIOP derived, 2-year averaged $PM_{2.5}$ concentrations are explored as functions of aerosol vertical distribution, $PM_{2.5}$ to PM_{10} ratio, RH, aerosol type, and cloud presence above. Spatial sampling related biases as well as prognostic errors are also studied.

3.2.1 Prognostic errors in $C_{m2.5}$

As a first step for the uncertainty analysis, we estimated the prognostic error of 2-year averaged $PM_{2.5\_CALIOP}$. Figure 5 shows the root-mean-square error (RMSE) of CALIOP-based $PM_{2.5}$ concentrations against those from EPA stations as a function of CALIOP-based $PM_{2.5}$ for the 2008-2009 period over the CONUS. RMSEs were computed for five equally sampled bins, determined from a cumulative histogram analysis. Each point in Fig. 5, from left to right, represents the RMSE and mean $PM_{2.5}$ concentration derived from CALIOP for 0-20%, 20-40%, 40-60%, 60-80%, and 80-100% cumulative frequencies. A mean combined daytime and nighttime RMSE of ~4 μg m^{-3} is found, with a mean value slightly greater for nighttime (~4.3 μg m^{-3}) than daytime (~3.7 μg m^{-3}). While most bins exhibit larger nighttime RMSEs, daytime RMSEs are larger for the greatest mean CALIOP-derived $PM_{2.5}$ concentrations.

Deleted: RMSEs were computed in intervals of 5 μg m^{-3} from 0 to 25 μg m^{-3}, with no computations greater than 25 μg m^{-3} performed due to very few data points above this $PM_{2.5}$ concentration level.

Deleted: minimum error

Deleted: with

Deleted: generally larger RMSEs for nighttime below 15 μg m^{-3}, and larger RMSEs for daytime above 15 μg m^{-3}. However, mean RMSEs (i.e., computed from the RMSEs shown in Fig. 5) are similar for both datasets, ~4.5 μg m^{-3} for daytime and ~4.0 μg m^{-3} for nighttime. Also, note that while the absolute error for daytime is largest at high $PM_{2.5}$ concentrations, relative errors are similar (e.g., 3 μg m^{-3}/10 μgm^{-3} or 30% for the 5-10 μg m^{-3} bin, versus 7 μg m^{-3}/25 μg m^{-3} or 28% for the 20-25 μg m^{-3} bin). For context, the number of samples per bin are also plotted (as X symbols) in Fig. 5. Data sample sizes are smallest (largest) for the lowest/highest range (mid-range) $PM_{2.5}$ bins. ¶

3.2.2 Surface layer height sensitivity study

A sensitivity study was conducted for which $PM_{2.5}$ was derived from near-surface CALIOP aerosol extinction by varying the height of the surface layer in increments of 100 m from the ground to 1000 m. Note that the surface layer (0-100 m) is included for this sensitivity study only. The statistical results of this analysis, for both daytime and nighttime conditions, are shown in Table 2. Four statistical parameters were computed, consisting of r^2, slope from Deming regression, mean bias (CALIOP – EPA) of $PM_{2.5}$, and percent error change in derived $PM_{2.5}$, defined as: ((mean_new_$PM_{2.5}$ – mean_original_$PM_{2.5}$)/mean_original_$PM_{2.5}$)*100. For context, the bottom row of Table 2 shows the results from the original analysis. In terms of r^2 and slope, optimal values peak at different surface layer heights between daytime and nighttime. For example, for daytime, the largest correlations are found for the 0-600 m and 0-700 m layers, while

Deleted: , slope

for nighttime these are found for the 0-300 m and 0-400 m layers. However, the 0-300 m layer exhibits the lowest mean bias for the daytime analysis, and the 100-1000 m layer exhibits the lowest mean bias for the nighttime analysis. Overall, marginal changes are found for varying the height of the surface layer. Yet the largest mean bias is found for the 0-100 m layer, indicating the need for excluding the 0-100 m layer in the analysis.

3.2.3 RH sensitivity study

Profiles of RH were taken from the MERRA-2 reanalysis product, as these collocated data are provided in the CALIPSO L2_05kmAPro product. However, biases may exist in this RH dataset. Thus, we examined the impact of varying the RH values by +/- 10% on the CALIOP-derived $PM_{2.5}$ concentrations. For both daytime and nighttime analyses, no significant differences in the r^2 and slope values were found. However, a +15% change in the mean derived $PM_{2.5}$ values was found by decreasing the RH values by 10%, while a -15% change in the mean derived $PM_{2.5}$ values was found by increasing the RH values by 10%.

3.2.4 $PM_{2.5}$ to PM_{10} ratio sensitivity study

Another source of uncertainty in this study is the $PM_{2.5}/PM_{10}$ ratio. Using surface-based $PM_{2.5}$ and PM_{10} data from those EPA stations over the CONUS for 2008-2009 with concurrent $PM_{2.5}$ and PM_{10} daily data available (i.e., 409 stations), we computed the mean $PM_{2.5}/PM_{10}$ ratio, and its corresponding standard deviation. The mean ratio was 0.56 with a standard deviation of 0.32. It is interesting to note that the mean $PM_{2.5}/PM_{10}$ ratio estimated from two years of surface observations over the CONUS is close to 0.6 (the number used in this study), as reported by Kaku et al. (2018). We also tested the sensitivity of the derived $PM_{2.5}$ concentrations as a function of $PM_{2.5}/PM_{10}$ ratio for two scenarios: ±1 standard deviation of the mean (Table 3). In general, a ±50

Deleted: (100-1000 m layer)

Deleted: (

Deleted:)

Formatted: Indent: First line: 0.5"

Deleted:

Deleted: However, a +15% (-15%) change in the mean derived $PM_{2.5}$ values was found by decreasing (increasing) the RH values by 10%.

Formatted: Subscript

Formatted: Subscript

% to 60 % change is found with the variation of the $PM_{2.5}/PM_{10}$ ratio at the range of ±1 standard deviation of the mean. As suggested from Table 3, the lowest mean daytime bias is found for a ratio of 0.6, and for nighttime the lowest mean bias occurs using a ratio of 0.88.

Deleted: , the optimal slope is found using a ratio of +1 standard deviation of the mean for both daytime and nighttime.

Deleted: T

3.2.5 Sampling-related biases

As mentioned in the introduction section, a sampling bias, due to the very small footprint size and ~16 day repeat cycle of CALIOP, can exist when using CALIOP observations for $PM_{2.5}$ estimates (Zhang and Reid, 2009). This sampling-induced bias is investigated from a 2-year mean perspective by comparing histograms of $PM_{2.5\_EPA}$ and $C_{m2.5}$ concentrations as shown in Fig. 6. To generate Fig. 6, all available daily EPA $PM_{2.5}$ are used to represent the "true" 2-year mean spectrum of $PM_{2.5}$ concentrations over the EPA sites. The aerosol extinction data spatially collocated to the EPA sites (Sect. 3.1), but not temporally collocated, are used for estimating the 2-year mean spectrum of $PM_{2.5}$ concentrations as derived from CALIOP observations. To be consistent with the previous analysis, only cloud-free CALIOP profiles are considered. The $PM_{2.5\_EPA}$ concentrations peak at ~10 μg m^{-3} (standard deviation of ~3 μg m^{-3}), and CALIOP-derived $PM_{2.5}$ peaks at ~9 μg m^{-3} (daytime; standard deviation of ~4 μg m^{-3}) and ~7 μg m^{-3} (nighttime; standard deviation of ~2 μg m^{-3}). The distribution shifts towards smaller concentrations for CALIOP, more so for nighttime than daytime (possibly due to CALIOP daytime versus nighttime detection differences).

Deleted: 9

Deleted: 5

Still, Fig. 6 may reflect the diurnal difference in $PM_{2.5}$ concentrations as well as the retrieval bias in $C_{m2.5}$ values. Thus, we have re-performed the exercise shown in Fig. 6 using spatially and temporally collocated $PM_{2.5\_EPA}$ and $C_{m2.5}$ data as shown in Fig. 7. To construct Fig. 7, $PM_{2.5\_EPA}$ and $C_{m2.5}$ data are collocated following the steps mentioned in Sect. 3.1, with CALIOP

and EPA $PM_{2.5}$ representing 2-year mean values for each EPA station. Again, only cloud-free CALIOP profiles are considered for this analysis. As shown in Fig. 7a, the $PM_{2.5\_EPA}$ concentrations peak at ~12 μg m^{-3} (standard deviation of ~4 μg m^{-3}), and daytime $C_{m2.5}$ peaks at ~10 μg m^{-3} (standard deviation of ~4 μg m^{-3}). In comparison, with the use of collocated nighttime $C_{m2.5}$ and $PM_{2.5\_EPA}$ data as shown in Fig. 7b, the peak $PM_{2.5\_EPA}$ value is about 5 μg m^{-3} higher than the peak $C_{m2.5}$ value (with similar standard deviations as found in the analyses of Fig. 7a). Considering both Figs. 6 and 7, it is likely that the temporal sampling bias seen in Fig. 6 is at least in part due to retrieval bias as well as the difference in $PM_{2.5}$ concentrations during daytime and nighttime.

Deleted: 7

Deleted: 6

Deleted: 2

3.2.6 CALIOP AOD analysis

Most past studies focused on the use of column AODs as proxies for surface $PM_{2.5}$ (e.g., Liu et al., 2005; Hoff and Christopher, 2009; van Donkelaar et al., 2015). Therefore, it is interesting to investigate whether near surface CALIOP extinction values can be used as a better physical quantity to estimate surface $PM_{2.5}$ in comparing with column-integrated CALIOP AOD. To achieve this goal, we have compared CALIOP column AOD and $PM_{2.5}$ from EPA stations, as shown in Fig. 8. Similar to the scatterplots of Fig. 4, each point represents a two-year mean for each EPA site, and was created from a dataset following the same spatial/temporal collocation as described above. As shown in Fig. 9, r^2 values of 0.04 and 0.13 are found using CALIOP daytime and nighttime AOD data, respectively, similar to the MODIS-based analysis shown in Fig. 1. This is expected, as elevated aerosol layers will negatively impact the relationship between surface $PM_{2.5}$ and column AOD. The derivation of surface $PM_{2.5}$ from near surface CALIOP extinction, as demonstrated from this study however, provides a much better spatial matching between the

quantities being compared, with potential error terms that can be well quantified and minimized in later studies.

3.2.7 Cloud flag sensitivity study

For most of this paper, a strict cloud screening process is implemented, during which no clouds are allowed in the entire CALIOP profile. However, in contrast to passive sensor capabilities (e.g., MODIS), near-surface aerosol extinction coefficients can be readily retrieved from CALIOP profiles even when there are transparent cloud layers above. Therefore, we conducted an additional analysis for which no cloud flag was set (i.e., all-sky conditions). Results are shown in scatterplot form in Fig. 9, in a similar manner as Figs. 3e and f, with an additional 97 points for the daytime analysis and 156 points for the nighttime analysis. Comparing the all-sky results with those of Figs. 3e, and f (cloud-free conditions), the r^2 values are similar. This is also true in terms of mean bias, with similar values of 0.70 μg m^{-3} found for daytime, and -2.68 μg m^{-3} for nighttime, all-sky scenarios. This indicates that our method performs reasonably well from an all-sky perspective. However, we note that restricting the analysis to solely those cases that are cloudy (not shown), the method does not perform as well. For example, the r^2 value decreases by 71% for the daytime analysis compared to the cloud-free results (Fig. 3e). The corresponding nighttime r^2 value decreases by 90%. This is expected, as any errors made in estimating the optical depths of the overlying clouds will propagate (as biases) into the extinction retrievals for the underlying aerosols.

3.2.8 Aerosol type analysis

Deleted: with an additional 97 (156) points for the daytime (nighttime) analyses

Deleted: .

Deleted: (-2.68)

Deleted:

Deleted: (nighttime)

Deleted:

Deleted: s

Deleted: (

Deleted:)

Deleted: and the slope values decrease by 21% (75%) for the daytime (nighttime) analyses, compared to the cloud-free results (Figs. 3e and f).

Also, for this study, we assume that the primary aerosol type over the CONUS is pollution (i.e., sulfate) aerosol, which is generally composed of smaller (fine mode) particles that tend to exhibit mass extinction efficiencies ~4 $m^2 g^{-1}$. However, even after implementing our dust-free restriction, the study region can also be contaminated with non-pollution aerosols, which can have a larger particle size and exhibit lower mass extinction efficiencies (e.g., Hess et al., 1998; Malm and Hand, 2007; Lynch et al., 2016). The use of $PM_{2.5}$ versus PM_{10} somewhat mitigates this size dependency, but nevertheless coarse mode dust or sea salt can dominate $PM_{2.5}$ mass values (e.g., Atwood et al., 2013).

Thus, in this section, the impact of aerosol types to the derived $PM_{2.5}$ concentrations was explored by varying the mass scattering and absorption efficiencies and gamma values associated with each aerosol type. The three aerosol types chosen for this sensitivity study were dust, sea salt, and smoke, based upon Lynch et al. (2016). The mass scattering and absorption values for dust and sea salt were interpolated to 0.532 µm from values at 0.450 µm and 0.550 µm from OPAC (as was done for the sulfate case; Hess et al., 1998). For smoke, these values were interpolated to 0.532 µm from values at 0.440 µm and 0.670 µm as provided by Reid et al. (2005) for smoke cases over the US and Canada. The gamma values were taken from Lynch et al. (2016) and the references within. These values, as well as the results from this sensitivity study, are shown in Table 4. If we assume all aerosols within the study region are smoke aerosols, no major changes in the retrieved CALIOP $PM_{2.5}$ values are found. However, significant uncertainties on the order of ~200% are found if sea salt, or ~800% if dust, aerosol mass scattering/absorption efficiencies and gamma values are used instead. Clearly, this study suggests that accurate aerosol typing is necessary for future applications of CALIOP observations for surface $PM_{2.5}$ estimations.

Deleted: & organic

Deleted: (~800%)

Deleted: (

Deleted:)

3.3 E-folding correlation length for $PM_{2.5}$ concentrations over the CONUS

As a last study, we also estimated the spatial e-folding correlation length for $PM_{2.5}$ concentrations over the CONUS. This provides us an estimation of the correlation between a CALIOP-derived and actual $PM_{2.5}$ concentration for a given location as a function of distance between the CALIOP observation and the given location. To accomplish this, all EPA stations over the CONUS with at least 50 days of daily data available for the 2008-2009 period were first determined. Next, the distances between each pair of these EPA stations, and their corresponding correlation of daily $PM_{2.5}$ concentrations, were computed. Results are shown in Fig. 10 as a scatterplot, with individual points in gray and the black curve representing the exponential fit to the data. A decrease in $PM_{2.5}$ correlation with distance between EPA stations is found, and the e-folding length in correlation (e.g., correlation reduced to 1/e, or 0.37) is ~600 km (from an AOD standpoint, this value is 40-400 km, as suggested by Anderson et al., 2003).

Also included in Fig. 10 are results from a corresponding regional analysis, with the red and blue lines showing bin averages (10 km) for the Western and Eastern CONUS, respectively (regions partitioned by the -97° longitude line). The e-folding length is ~300 km for the Western CONUS, and ~700 km for the Eastern CONUS, indicating a much shorter correlation length for pollution over the Western CONUS, possibly due to a more complex terrain such as mountains. Overall, these $PM_{2.5}$ e-folding lengths suggest that CALIOP-derived $PM_{2.5}$ concentrations could still have some representative skill within a few hundred kilometers of a given location.

4 Conclusions

In this paper, we have demonstrated a new bulk-mass-modeling method for retrieving surface particulate matter (PM) with particle diameters smaller than 2.5 μm ($PM_{2.5}$) using

Deleted: 2.9

Deleted: (

Deleted:)

Deleted: Western (

Deleted:)

Deleted: sizes

observations acquired by the NASA Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument from 2008-2009. For the purposes of demonstrating this concept, only regionally-averaged parameters, such as mass scattering and absorption coefficients, and $PM_{2.5}$ to PM_{10} (PM with particle diameters smaller than 10 μm) conversion ratio, are used. Also, we assume the dominant type of aerosols over the study region is pollution aerosols (supported by the occurrence frequencies of aerosol types determined by the CALIOP algorithms), and exclude aerosol extinction range bins classified as dust from the analysis. Even with the highly-averaged parameters, the results from this paper are rather promising and demonstrate a potential for monitoring PM pollution using active-based lidar observations. Specifically, the primary results of this study are as follows:

1. CALIOP-derived $PM_{2.5}$ concentrations of ~10-12.5 μg m^{-3} are found over the eastern contiguous United States (CONUS), with lower values of ~2.5-5 μg m^{-3} over the central CONUS. $PM_{2.5}$ values of ~10-20 μg m^{-3} are found over the west coast of the CONUS, primarily California. The spatial distribution of 2-year mean $PM_{2.5}$ concentrations derived from near surface CALIOP aerosol data compares well to the spatial distribution of *in situ* $PM_{2.5}$ measurements collected at the ground-based stations of the U.S. Environmental Protection Agency (EPA). The use of nighttime CALIOP extinction to derive $PM_{2.5}$ results in a higher correlation (r^2 = 0.48; mean bias = -3.3 μgm^{-3}) with EPA $PM_{2.5}$ than daytime CALIOP extinction data (r^2 = 0.21; mean bias = -0.40 μgm^{-3}).
2. Correlations between CALIOP aerosol optical depth (AOD) and EPA $PM_{2.5}$ are much lower (r^2 values of 0.04 and 0.13, for daytime and nighttime CALIOP AOD data, respectively) than those obtained from derived $PM_{2.5}$ using near-surface CALIOP aerosol extinction. A similar correlation is also found between Moderate Resolution Imaging

Deleted: sizes

Spectroradiometer (MODIS) AOD and EPA $PM_{2.5}$ from two-year (2008-2009) means. This suggests that CALIOP extinction may be used as a better parameter for estimating $PM_{2.5}$ concentrations from a 2-year mean perspective. Also, the algorithm proposed in this study is essentially a semi-physical-based method, and thus the retrieval process can be improved, upon a careful study of the physical parameters used in the process.

3. Spatial and temporal sampling biases, as well as a retrieval bias, are found. Also, several sensitivity studies were conducted, including surface layer height, cloud flag, $PM_{2.5}/PM_{10}$ ratio, relative humidity, and aerosol type. The sensitivity studies highlight the need for accurate aerosol typing for estimating $PM_{2.5}$ concentrations using CALIOP observations.
4. Using surface-based $PM_{2.5}$ at EPA stations alone, the e-folding correlation length for $PM_{2.5}$ concentrations was found to be about 600 km for the CONUS. A regional analysis yielded values of ~300 km and ~700 km for the Western and Eastern CONUS, respectively. Thus, while limited in spatial sampling, measurements from CALIOP may still be used for estimating $PM_{2.5}$ concentrations over the CONUS.

As noted earlier, CALIOP observations are still rather sparse, and concerns related to reported CALIOP aerosol extinction values also exist, such as solar and surface contamination and the "retrieval fill value" issue (e.g., Toth et al., 2018). Yet, the future High Spectral Resolution Lidar (HSRL) instrument on board the Earth Clouds, Aerosol, and Radiation Explorer (EarthCARE) satellite (Illingworth et al., 2015), as well as forthcoming space-based lidar missions in response to the 2017 Decadal Survey, offer opportunities to further explore aerosol extinction - based PM concentrations. Ultimately the results from this study show that the combined use of several lidar instruments for monitoring regional and global PM pollution is potentially achievable.

Acknowledgements

This research was funded with the support of the NASA Earth and Space Science Fellowship program (NNX16A066H). Author JZ acknowledges the support from NASA grant NNX17AG52G. CALIPSO data were obtained from the NASA Langley Research Center Atmospheric Science Data Center (eos-web.larc.nasa.gov). MODIS data were obtained from NASA Goddard Space Flight Center (ladsweb.nascom.nasa.gov). The $PM_{2.5}$ data were obtained from the EPA AQS site (https://aqs.epa.gov/aqsweb/airdata/download_files.html).

References

Anderson, T.L., Charlson, R. J., Winker, D.M., Ogren, J.A., and Holmén, K.: Mesoscale Variations of Tropospheric Aerosols, J. Atmos. Sci., 60, 119–136, https://doi.org/10.1175/1520-0469(2003)060<0119:MVOTA>2.0.CO;2, 2003.

Atwood, S. A., J. S. Reid, S. M. Kreidenweis, S. S. Cliff, Y. Zhao, N.-H. Lin, S.-C.Tsay, Y.-C. Chu, and Westphal, D.L.: Size resolved measurements of springtime aerosol particles over the northern South China Sea, Atmospheric Environment, 78, 134-143, https://doi.org/10.1016/j.atmosenv.2012.11.024, 2013.

Campbell, J. R., Tackett, J. L., Reid, J. S., Zhang, J., Curtis, C. A., Hyer, E. J., Sessions, W. R., Westphal, D. L., Prospero, J. M., Welton, E. J., Omar, A. H., Vaughan, M. A., and Winker, D. M.: Evaluating nighttime CALIOP 0.532 μm aerosol optical depth and extinction coefficient retrievals, Atmos. Meas. Tech., 5, 2143-2160, https://doi.org/10.5194/amt-5-2143-2012, 2012.

Charlson, R. J., Ahlquist, N. C., and Horvath, H.: On the generality of correlation of atmospheric aerosol mass concentration and light scatter, Atmospheric Environment, 2(5), 455-464, https://doi.org/10.1016/0004-6981(68)90039-5, 1968.

Chew, B. N., Campbell, J. R., Hyer, E. J., Salinas, S. V., Reid, J. S., Welton, E. J., Holben, B.N. and Liew, S. C.: Relationship between aerosol optical depth and particulate matter over Singapore: Effects of aerosol vertical distributions, Aerosol and Air Quality Research, 16, 2818-2830, https://doi.org/10.4209/aaqr.2015.07.0457, 2016.

Chow, J. C., Watson, J. G., Park, K., Robinson, N. F., Lowenthal, D. H., Park, K., and Magliano, K. A.: Comparison of particle light scattering and fine particulate matter mass in central California, Journal of the Air & Waste Management Association, 56(4), 398-410, https://doi.org/10.1080/10473289.2006.10464515, 2006.

Chung, A., Chang, D. P., Kleeman, M. J., Perry, K. D., Cahill, T. A., Dutcher, D., ... & Stroud, K: Comparison of real-time instruments used to monitor airborne particulate matter, Journal of the Air & Waste Management Association, 51(1), 109-120, 2001.

Colarco, P. R., Kahn, R. A., Remer, L. A., and Levy, R. C.: Impact of satellite viewing-swath width on global and regional aerosol optical thickness statistics and trends, Atmos. Meas. Tech., 7, 2313-2335, https://doi.org/10.5194/amt-7-2313-2014, 2014.

Deming, W.E.: Statistical Adjustment of Data, Wiley: New York, 1943.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., ... and Schwartz, J. D.: Air pollution and mortality in the Medicare population, New England Journal of Medicine, 376(26), 2513-2522, doi: 10.1056/NEJMoa170274, 2017.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L.: Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge, Atmospheric Environment, 37(9-10), 1277-1292, 2003.

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., and Hoff, R. M.: Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality, Atmos. Environ., 38, 2495–2509, https://doi.org/10.1016/j.atmosenv.2004.01.039, 2004.

Federal Register: National ambient air quality standards for particulate matter. Final Rule Federal Register/vol. 62, no. 138/18 July 1997/Final Rule, 40 CFR Part 50, 1997.

Glantz, P., Kokhanovsky, A., von Hoyningen-Huene, W., and Johansson, C.: Estimating PM2.5 over southern Sweden using space-borne optical measurements, Atmospheric Environment, 43(36), 5838-5846, https://doi.org/10.1016/j.atmosenv.2009.05.017, 2009.

Gong, W., Huang, Y., Zhang, T., Zhu, Z., Ji, Y., and Xiang, H.: Impact and Suggestion of

Column-to-Surface Vertical Correction Scheme on the Relationship between Satellite AOD and Ground-Level PM2.5 in China, Remote Sensing, 9(10), 1038, https://doi.org/10.3390/rs9101038, 2017.

Greenstone, M.: The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures, Journal of political economy, 110(6), 1175-1219, https://doi.org/10.1086/342808, 2002.

Hand, J. L., and Malm, W. C.: Review of aerosol mass scattering efficiencies from ground-based measurements since 1990, Journal of Geophysical Research: Atmospheres, 112(D16), https://doi.org/10.1029/2007JD008484, 2007.

Hand, J. L., Schichtel, B. A., Malm, W. C., and Frank, N. H.: Spatial and temporal trends in PM2.5 organic and elemental carbon across the United States, Advances in Meteorology, http://dx.doi.org/10.1155/2013/367674, 2013.

Hänel, G.: The properties of atmospheric aerosol particles as functions of the relative humidity at thermodynamic equilibrium with the surrounding moist air, Advances in geophysics, 19, 73-188, https://doi.org/10.1016/S0065-2687(08)60142-9, 1976.

Hess, M., Koepke, P., and Schult, I.: Optical properties of aerosols and clouds: The software package OPAC, Bulletin of the American meteorological society, 79(5), 831-844, https://doi.org/10.1175/1520-0477(1998)079%3C0831:OPOAAC%3E2.0.CO;2, 1998.

Hoff, Raymond M., and Christopher, Sundar A.: Remote sensing of particulate pollution from space: have we reached the promised land?, Journal of the Air & Waste Management Association, 59.6 (2009): 645-675, https://doi.org/10.3155/1047-3289.59.6.645, 2009.

Huang, X. H., Bian, Q., Ng, W. M., Louie, P. K., and Yu, J. Z: Characterization of PM2.5 major

components and source investigation in suburban Hong Kong: a one year monitoring study, Aerosol Air Qual. Res, 14(1), 237-250, 2014.

Hunt, W. H., Winker, D. M., Vaughan, M. A., Powell, K. A., Lucker, P. L., and Weimer, C.: CALIPSO lidar description and performance assessment, Journal of Atmospheric and Oceanic Technology, 26(7), 1214-1228, https://doi.org/10.1175/2009JTECHA1223.1, 2009.

Illingworth, Anthony J., et al.: The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation, Bulletin of the American Meteorological Society, 96.8, 1311-1332, https://doi.org/10.1175/BAMS-D-12-00227.1, 2015.

Kaku, K. C., Reid, J. S., Hand, J. L., Edgerton, E. S., Holben, B. N., Zhang, J., and Holz, R. E.: Assessing the challenges of surface-level aerosol mass estimates from remote sensing during the SEAC4RS and SEARCH campaigns: Baseline surface observations and remote sensing in the southeastern United States, Journal of Geophysical Research: Atmospheres, 123, 7530–7562, https://doi.org/10.1029/2017JD028074, 2018.

Kiss, G., Imre, K., Molnár, Á., and Gelencsér, A.: Bias caused by water adsorption in hourly PM measurements, Atmos. Meas. Tech., 10, 2477-2484, https://doi.org/10.5194/amt-10-2477-2017, 2017.

Kittaka, C., Winker, D. M., Vaughan, M. A., Omar, A., & Remer, L. A.: Intercomparison of column aerosol optical depths from CALIPSO and MODIS-Aqua, Atmospheric Measurement Techniques, 4(2), 131, https://doi.org/10.5194/amt-4-131-2011, 2011.

Kumar, N., Chu, A., and Foster, A.: An empirical relationship between PM2.5 and aerosol optical

depth in Delhi Metropolitan, Atmos. Environ., 41, 4492–4503, https://doi.org/10.1016/j.atmosenv.2007.01.046, 2007.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N.C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Li, J., Carlson, B. E., and Lacis, A. A.: How well do satellite AOD observations represent the spatial and temporal variability of PM2.5 concentration for the United States?, Atmospheric environment, 102, 260-273, https://doi.org/10.1016/j.atmosenv.2014.12.010, 2015.

Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P.: Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison between MISR and MODIS, Remote sensing of Environment, 107(1-2), 33-44, https://doi.org/10.1016/j.rse.2006.05.022, 2007.

Liu, Y., Park, R. J., Jacob, D. J., Li, Q., Kilaru, V., and Sarnat, J. A.: Mapping annual mean ground-level PM2. 5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States, Journal of Geophysical Research: Atmospheres, 109 (D22), https://doi.org/10.1029/2004JD005025, 2004.

Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., and Koutrakis, P.: Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing, Environmental science & technology, 39(9), 3269-3278, doi: 10.1021/es049352m, 2005.

Liou, Kuo-Nan.: An introduction to atmospheric radiation, Vol. 84. Academic press, 2002.

Lynch P., and coauthors: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for

atmospheric and climate sciences, Geosci. Model Dev., 9, 1489-1522, doi:10.5194/gmd-9-1489-2016, 2016.

Malm, W.C. and Hand, J.L.: An examination of the physical and optical properties of aerosols collected in the IMPROVE program, Atmospheric Environment, 41(16), pp.3407-3427, https://doi.org/10.1016/j.atmosenv.2006.12.012, 2007.

Nessler, R., Weingartner, E., and Baltensperger, U. (2005). Effect of humidity on aerosol light absorption and its implications for extinction and the single scattering albedo illustrated for a site in the lower free troposphere, *Journal of Aerosol Science*, *36*(8), 958-972.

Omar, A. H., Winker, D. M., Tackett, J. L., Giles, D. M., Kar, J., Liu, Z., Vaughan, M. A., Powell, K. A., and Trepte, C. R.: CALIOP and AERONET aerosol optical depth comparisons: One size fits none, J. Geophys. Res. Atmos., 118, 4748–4766, doi:10.1002/jgrd.50330, 2013.

Patashnick, H., Rupprecht, G., Ambs, J. L., and Meyer, M. B.: Development of a reference standard for particulate matter mass in ambient air, Aerosol Science & Technology, 34(1), 42-45, 2001.

Reid, J. S., Eck, T. F., Christopher, S. A., Koppmann, R., Dubovik, O., Eleuterio, D. P., Holben, B. N., Reid, E. A., and Zhang, J.: A review of biomass burning emissions part III: intensive optical properties of biomass burning particles, Atmos. Chem. Phys., 5, 827-849, https://doi.org/10.5194/acp-5-827-2005, 2005.

Reid, J. S. et al.: Skill of Operational Aerosol Forecast Models in Predicting Aerosol Events and Trends of the Eastern United States, A11B-001, AGU Fall meeting, San Francisco, 12-16 Dec, 2016.

Reid, J. S., Kuehn, R. E., Holz, R. E., Eloranta, E. W., Kaku, K. C., Kuang, S., ... and Atwood,

S.A.: Ground-based High Spectral Resolution Lidar observation of aerosol vertical distribution in the summertime Southeast United States, Journal of Geophysical Research: Atmospheres, 122(5), 2970-3004, https://doi.org/10.1002/2016JD025798, 2017.

Sessions, W. R., Reid, J. S., Benedetti, A., Colarco, P. R., da Silva, A., Lu, S., Sekiyama, T., Tanaka, T. Y., Baldasano, J. M., Basart, S., Brooks, M. E., Eck, T. F., Iredell, M., Hansen, J. A., Jorba, O. C., Juang, H.-M. H., Lynch, P., Morcrette, J.-J., Moorthi, S., Mulcahy, J., Pradhan, Y., Razinger, M., Sampson, C. B., Wang, J., and Westphal, D. L.: Development towards a global operational aerosol consensus: basic climatological characteristics of the International Cooperative for Aerosol Prediction Multi-Model Ensemble (ICAP-MME), Atmos. Chem. Phys., 15, 335-362, https://doi.org/10.5194/acp-15-335-2015, 2015.

Silva, R. A., West, J. J., Zhang, Y., Anenberg, S. C., Lamarque, J. F., Shindell, D. T., Collins, W.J., Dalsoren, S., Faluvegi, G., Folberth, G. & Horowitz, L. W.: Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change, Environmental Research Letters, 8(3), 034005, doi:10.1088/1748-9326/8/3/034005, 2013.

Spagnolo, G. S.: Automatic instrument for aerosol samples using the beta-particle attenuation, Journal of aerosol science, 20(1), 19-27, 1989.

Toth, T. D., Campbell, J. R., Reid, J. S., Tackett, J. L., Vaughan, M. A., Zhang, J., and Marquis, J. W.: Minimum aerosol layer detection sensitivities and their subsequent impacts on aerosol optical thickness retrievals in CALIPSO level 2 data products, Atmos. Meas. Tech., 11, 499-514, https://doi.org/10.5194/amt-11-499-2018, 2018.

Toth, T. D., Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Shi, Y., and Westphal, D. L.: Impact

of data quality and surface-to-column representativeness on the $PM_{2.5}$ / satellite AOD relationship for the contiguous United States, Atmos. Chem. Phys., 14, 6049-6062, https://doi.org/10.5194/acp-14-6049-2014, 2014.

Toth, T. D., Zhang, J., Campbell, J. R., Reid, J. S., Shi, Y., Johnson, R. S., Smirnov, A., Vaughan, M.A. and Winker, D. M.: Investigating enhanced Aqua MODIS aerosol optical depth retrievals over the mid-to-high latitude Southern Oceans through intercomparison with co-located CALIOP, MAN, and AERONET data sets, Journal of Geophysical Research: Atmospheres, 118(10), 4700-4714, https://doi.org/10.1002/jgrd.50311, 2013.

Toth, T. D., Zhang, J., Campbell, J. R., Reid, J. S., and Vaughan, M. A.: Temporal variability of aerosol optical thickness vertical distribution observed from CALIOP, Journal of Geophysical Research: Atmospheres, 121(15), 9117-9139, https://doi.org/10.1002/2015JD024668, 2016.

Val Martin, M., Heald, C. L., Ford, B., Prenni, A. J., and Wiedinmyer, C.: A decadal satellite analysis of the origins and impacts of smoke in Colorado, Atmos. Chem. Phys., 13, 7429-7439, https://doi.org/10.5194/acp-13-7429-2013, 2013.

Van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J.: Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite Based Aerosol Optical Depth: Development and Application, Environ. Health Perspect., 118(6): 847–855, https://dx.doi.org/10.1289%2Fehp.0901623, 2010.

Van Donkelaar, A., Martin, R. V., Spurr, R. J., and Burnett, R. T.: High-resolution satellite-derived PM2.5 from optimal estimation and geographically weighted regression over North America, Environmental science & technology, 49.17 (2015): 10482-10491, doi: 10.1021/acs.est.5b02076, 2015.

Waggoner, A. P., and Weiss, R. E.: Comparison of fine particle mass concentration and light scattering extinction in ambient aerosol, Atmospheric Environment (1967), 14(5), 623-626, https://doi.org/10.1016/0004-6981(80)90098-0, 1980.

Wang, J., and Christopher, S. A.: Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: implications for air quality studies, Geophysical research letters, 30(21), https://doi.org/10.1029/2003GL018174, 2003.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophysical Research Letters, 34(19), https://doi.org/10.1029/2007GL030135, 2007.

Winker, D. M., Pelon, J., Coakley Jr, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., ... and Kubar, T. L.: The CALIPSO mission: A global 3D view of aerosols and clouds, Bulletin of the American Meteorological Society, 91(9), 1211-1230, https://doi.org/10.1175/2010BAMS3009.1, 2010.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, J. Atmos. Oceanic Technol., 26, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1, 2009.

Young, S. A., Vaughan, M. A., Kuehn, R. E., and Winker, D. M.: The retrieval of profiles of particulate extinction from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data: Uncertainty and error sensitivity analyses, Journal of Atmospheric and Oceanic Technology, 30(3), 395-428, https://doi.org/10.1175/2009JTECHA1281.1, 2013.

Zhang, J., Campbell, J.R., Hyer, E. J., Reid, J.S., Westphal, D. L., and Johnson, R. S.: Evaluating

the impact of multisensor data assimilation on a global aerosol particle transport model, J. Geophys. Res. Atmos., 119, 4674–4689, doi:10.1002/2013JD020975, 2014.

Zhang, J. and Reid, J.S.: An analysis of clear sky and contextual biases using an operational over ocean MODIS aerosol product, Geophysical Research Letters, 36, L15824, doi:10.1029/2009GL038723, 2009.

Figure and Table Captions

Figure 1. For 2008-2009, scatterplot of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and mean column AOD (550 nm) from collocated Collection 6 (C6) Aqua MODIS observations. The red line represents the Deming regression fit.

Figure 2. For 2008-2009 over the CONUS, (a) mean $PM_{2.5}$ concentration (μg m^{-3}) for those U.S. EPA stations with reported daily measurements, and (c) 1° x 1° average CALIOP-derived $PM_{2.5}$ concentrations for the 100–1000 m AGL atmospheric layer, using Equation 3, for combined daytime and nighttime conditions. Also shown are the pairwise $PM_{2.5}$ concentrations from (b) EPA daily measurements and (d) those derived from CALIOP (day and night combined), both averaged for each EPA station for the 2008-2009 period. For all four plots, values greater than 20 μgm^{-3} are colored red.

Figure 3. For 2008-2009 over the CONUS, 1° x 1° average CALIOP extinction, relative to the number of cloud-free 5 km CALIOP profiles in each 1° x 1° bin, for the 100 – 1000 m AGL atmospheric layer, for (a) daytime and (b) nighttime measurements. Also shown are the corresponding CALIOP-derived $PM_{2.5}$ concentrations, using Equation 3 for (c) daytime and (d) nighttime conditions. Values greater than 0.2 km^{-1} and 20 μg m^{-3} for (a, b) and (c, d), respectively, are colored red. Scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated near-surface CALIOP observations are shown in the bottom row, using (e) daytime and (f) nighttime CALIOP data. The red lines represent the Deming regression fits.

Figure 4. Scatterplot of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated near-surface CALIOP observations using combined daytime and nighttime CALIOP data. The red line represents the Deming regression fit.

Figure 5. Root-mean-square errors of CALIOP-derived $PM_{2.5}$ against EPA $PM_{2.5}$ as a function of CALIOP-derived $PM_{2.5}$, using both daytime (in red) and nighttime (in blue) CALIOP observations. The five bins are equally sampled based upon a cumulative histogram analysis, and each point from left to right represents the RMSE and mean $PM_{2.5}$ concentration derived from CALIOP for 0-20%, 20-40%, 40-60%, 60-80%, and 80-100% cumulative frequencies.

Figure 6. Two-year (2008-2009) histograms of mean $PM_{2.5}$ concentrations from the U.S. EPA (in black) and those derived from aerosol extinction using nighttime (in blue) and daytime (in red) CALIOP data. The U.S. EPA data shown are not collocated, while those derived using CALIOP are spatially (but not temporally) collocated, with EPA station observations.

Deleted: Root-mean-square errors of CALIOP-derived $PM_{2.5}$ against EPA $PM_{2.5}$ as a function of CALIOP-derived $PM_{2.5}$ (filled circles), and corresponding number of data samples per bin (X symbols), using both daytime (in red) and nighttime (in blue) CALIOP observations. ¶

Deleted: .

Figure 7. Two-year (2008-2009) histograms of mean $PM_{2.5}$ concentrations from the U.S. EPA and those derived from spatially and temporally collocated aerosol extinction using (a) daytime and (b) nighttime CALIOP data.

Figure 8. For 2008-2009, scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and mean column AOD from collocated CALIOP observations, using (a) daytime and (b) nighttime CALIOP data. The red lines represent the Deming regression fits.

Figure 9. For 2008-2009, scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated all-sky (including cloud-free and cloudy profiles) near-surface CALIOP observations, using (a) daytime and (b) nighttime CALIOP data. The red lines represent the Deming regression fits.

Figure 10. For 2008-2009 over the CONUS, scatterplot of distance (km) between any two U.S. EPA stations and the corresponding spatial correlation of $PM_{2.5}$ concentration between each pair of stations. The black curve represents the exponential fit to the data for the entire CONUS, and the red and blue dashed lines represent 10 km bin averages for the Western and Eastern CONUS, respectively.

Table 1. The parameters, and corresponding values, used to quality assure the CALIOP aerosol extinction profile.

Table 2. Statistical summary of a sensitivity analysis varying the height of the surface layer, including R^2, slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in $\mu g\ m^{-3}$, and percent error change in derived $PM_{2.5}$, defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.

Deleted: , slope

Table 3. Statistical summary of a sensitivity analysis varying the $PM_{2.5}$ to PM_{10} ratio used, including slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in $\mu g\ m^{-3}$, and percent error change in derived $PM_{2.5}$, defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.

Deleted: slope,

Table 4. Statistical summary of a sensitivity analysis varying the aerosol type assumed in the derivation of $PM_{2.5}$, including R^2, slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in µg m^{-3}, and percent error change in derived $PM_{2.5}$, defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.

Deleted: , slope

Appendix Figure 1. For 2008-2009 over the CONUS, for each 1° x 1° grid box, the number of days and CALIOP Level 2 5 km aerosol profiles used in the creation of the maps in Fig. 3 for (a, c) daytime and (b, d) nighttime measurements. Values greater than 400 profiles for (c, d) are colored red.

Figures

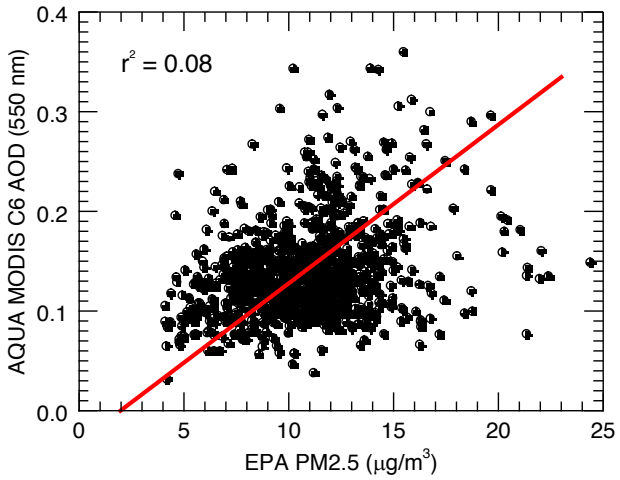


Figure 1. For 2008-2009, scatterplot of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and mean column AOD (550 nm) from collocated Collection 6 (C6) Aqua MODIS observations. The red line represents the Deming regression fit.

1462

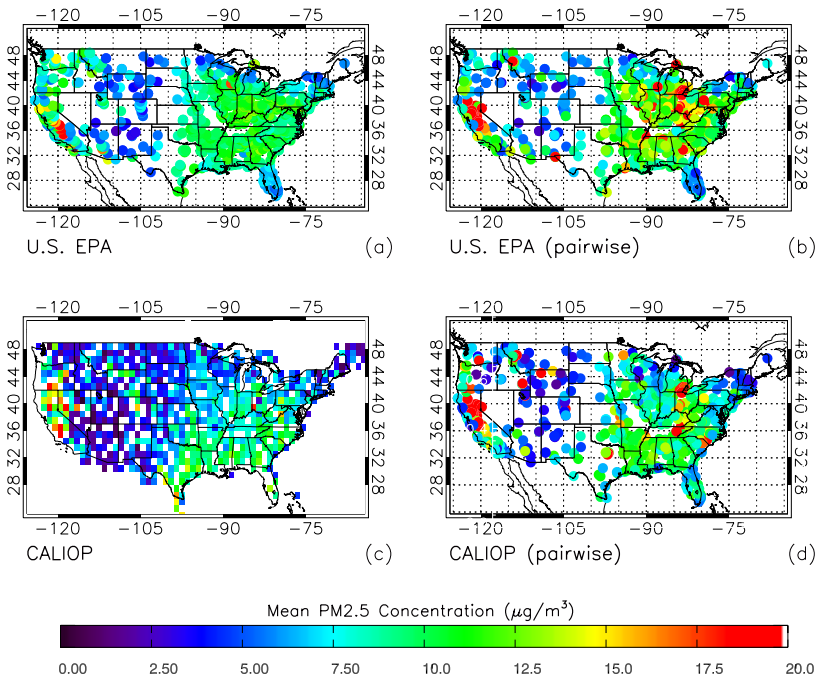


Figure 2. For 2008-2009 over the CONUS, (a) mean $PM_{2.5}$ concentration (μg m^{-3}) for those U.S. EPA stations with reported daily measurements, and (c) 1° x 1° average CALIOP-derived $PM_{2.5}$ concentrations for the 100–1000 m AGL atmospheric layer, using Equation 3, for combined daytime and nighttime conditions. Also shown are the pairwise $PM_{2.5}$ concentrations from (b) EPA daily measurements and (d) those derived from CALIOP (day and night combined), both averaged for each EPA station for the 2008-2009 period. For all four plots, values greater than 20 μg m^{-3} are colored red.

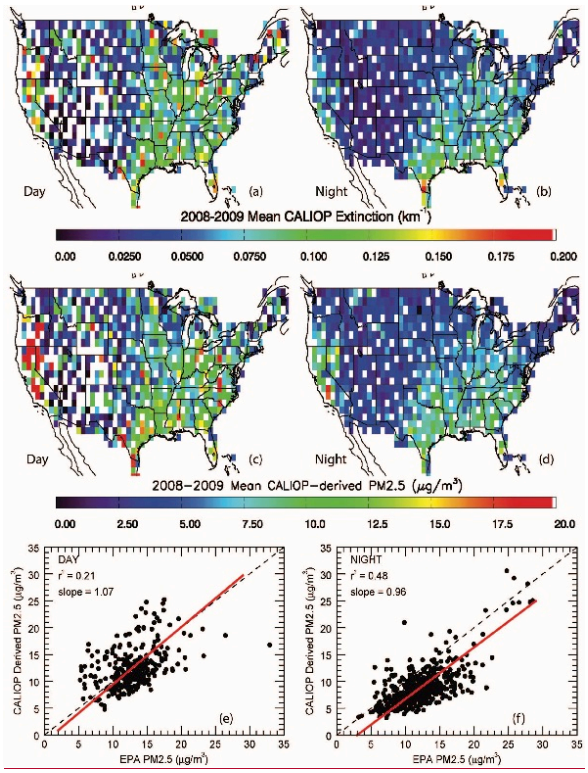


Figure 3. For 2008-2009 over the CONUS, 1° x 1° average CALIOP extinction, relative to the number of cloud-free L2_05kmAPro profiles in each 1° x 1° bin, for the 100 – 1000 m AGL atmospheric layer, for (a) daytime and (b) nighttime measurements. Also shown are the corresponding CALIOP-derived $PM_{2.5}$ concentrations, using Equation 3 for (c) daytime and (d) nighttime conditions. Values greater than 0.2 km^{-1} and 20 $\mu g\ m^{-3}$ for (a, b) and (c, d), respectively, are colored red. Scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated near-surface CALIOP observations are shown in the bottom row, using (e) daytime and (f) nighttime CALIOP data. The red lines represent the Deming regression fits.

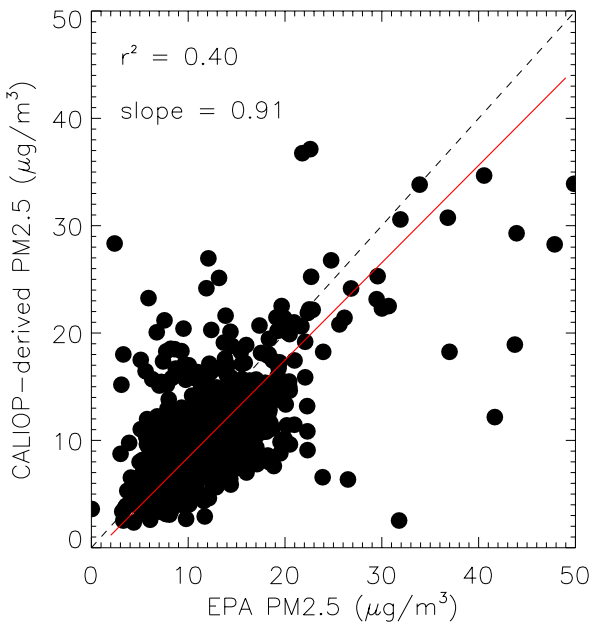


Figure 4. Scatterplot of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated near-surface CALIOP observations using combined daytime and nighttime CALIOP data. The red line represents the Deming regression fit.

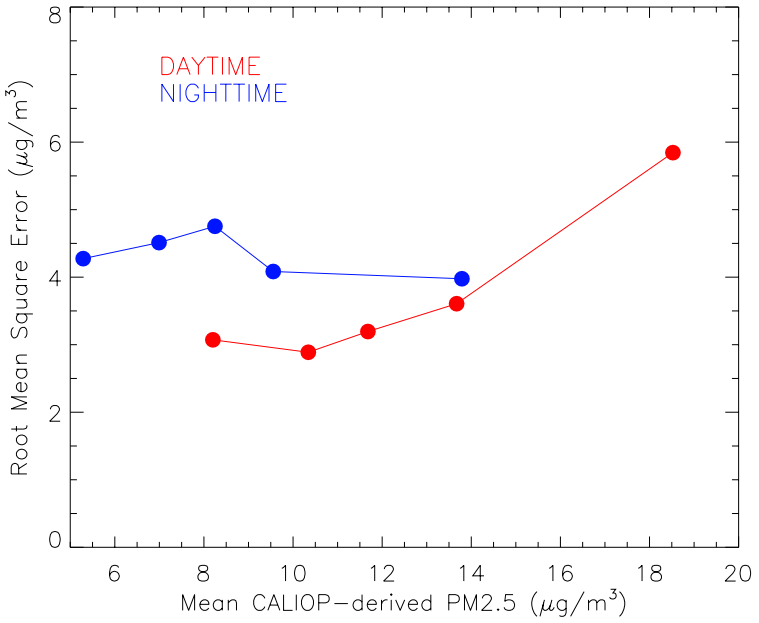


Figure 5. Root-mean-square errors of CALIOP-derived $PM_{2.5}$ against EPA $PM_{2.5}$ as a function of CALIOP-derived $PM_{2.5}$, using both daytime (in red) and nighttime (in blue) CALIOP observations. The five bins are equally sampled based upon a cumulative histogram analysis, and each point from left to right represents the RMSE and mean $PM_{2.5}$ concentration derived from CALIOP for 0-20%, 20-40%, 40-60%, 60-80%, and 80-100% cumulative frequencies.

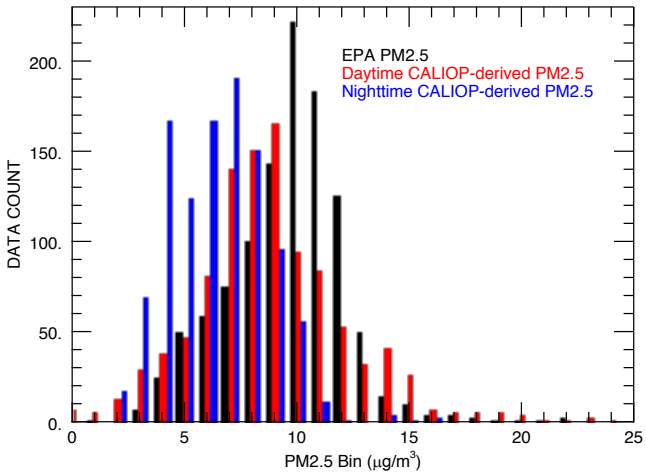


Figure 6. Two-year (2008-2009) histograms of mean $PM_{2.5}$ concentrations from the U.S. EPA (in black) and those derived from aerosol extinction using nighttime (in blue) and daytime (in red) CALIOP data. The U.S. EPA data shown are not collocated, while those derived using CALIOP are spatially (but not temporally) collocated, with EPA station observations.

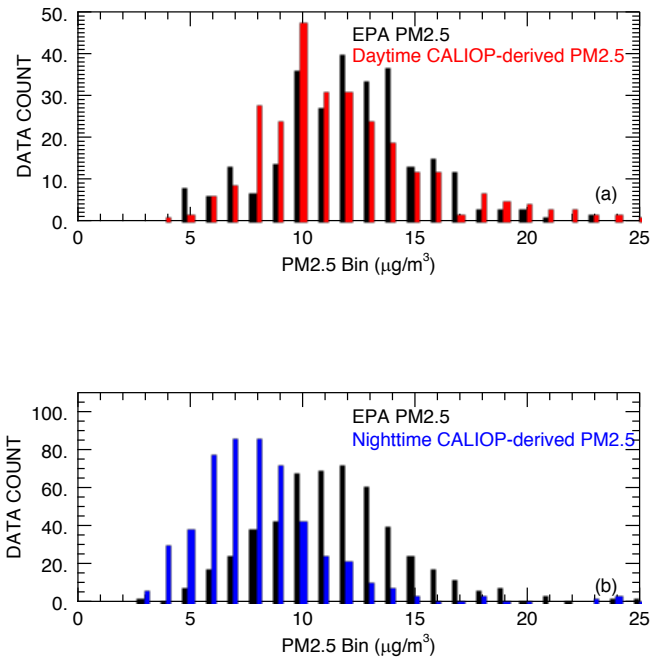


Figure 7. Two-year (2008-2009) histograms of mean $PM_{2.5}$ concentrations from the U.S. EPA and those derived from spatially and temporally collocated aerosol extinction using (a) daytime and (b) nighttime CALIOP data.

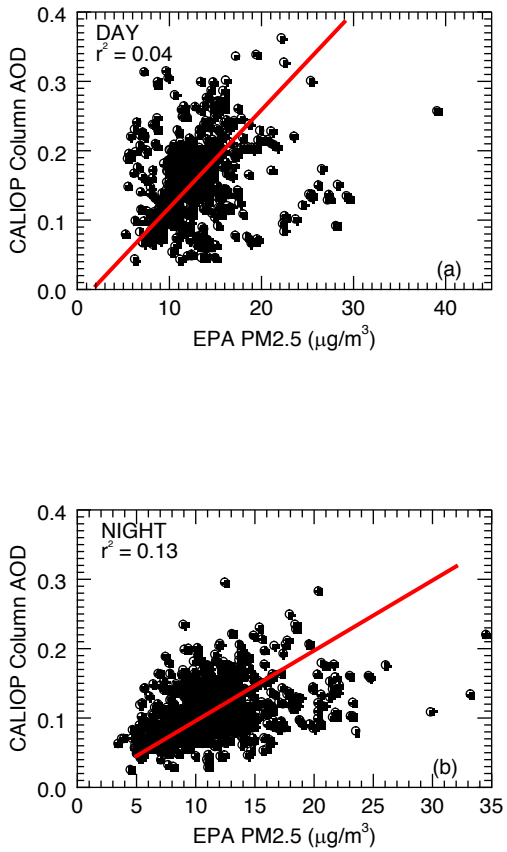


Figure 8. For 2008-2009, scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and mean column AOD from collocated CALIOP observations, using (a) daytime and (b) nighttime CALIOP data. The red lines represent the Deming regression fits.

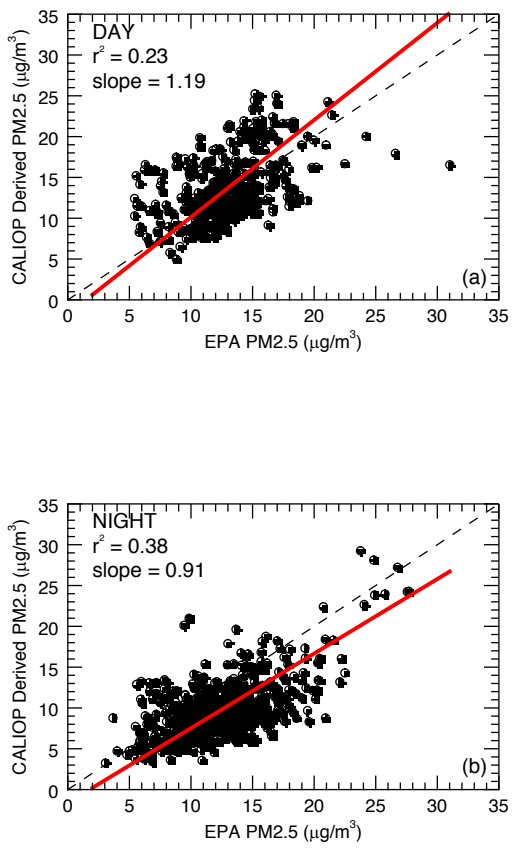


Figure 9. For 2008-2009, scatterplots of mean $PM_{2.5}$ concentration from ground-based U.S. EPA stations and those derived from collocated all-sky (including cloud-free and cloudy profiles) near-surface CALIOP observations, using (a) daytime and (b) nighttime CALIOP data. The red lines represent the Deming regression fits.

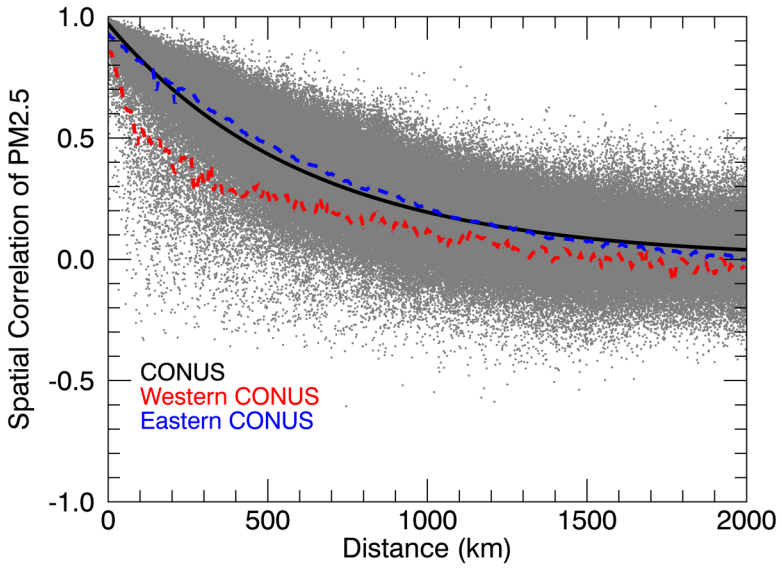


Figure 10. For 2008-2009 over the CONUS, scatterplot of distance (km) between any two U.S. EPA stations and the corresponding spatial correlation of $PM_{2.5}$ concentration between each pair of stations. The black curve represents the exponential fit to the data for the entire CONUS, and the red and blue dashed lines represent 10 km bin averages for the Western and Eastern CONUS, respectively.

Tables

Parameter	Values
Integrated_Attenuated_Backscatter_532	≤ 0.01 sr^{-1}
Extinction_Coefficient_532	≥ 0 and ≤ 1.25 km^{-1}
Extinction_QC_532	= 0, 1, 2, 16, or 18
CAD_Score	≥ -100 and ≤ -20
Extinction_Coefficient_Uncertainty_532	≤ 10 km^{-1}
Atmospheric_Volume_Description (Bits 1-3)	= 3
Atmospheric_Volume_Description (Bits 10-12)	$\neq 0$

Table 1. The parameters, and corresponding values, used to quality assure the CALIOP aerosol extinction profile.

Surface Layer (m)	Analysis (Day/Night)			
	R^2	Deming Slope	Mean Bias (CALIOP - EPA; μg m^{-3})	Error Change (%)
0-100	0.27/0.41	1.32/0.60	-2.67/-9.06	-13.71/-61.94
0-200	0.33/0.53	1.34/1.04	-0.52/-5.68	3.79/-23.58
0-300	0.35/0.54	1.32/1.11	-0.09/-4.70	7.24/-12.15
0-400	0.38/0.57	1.30/1.13	-0.13/-4.25	6.92/-6.46
0-500	0.35/0.52	1.26/1.06	-0.21/-4.04	5.70/-4.39
0-600	0.40/0.53	1.19/1.04	-0.46/-3.91	3.72/-2.15
0-700	0.44/0.46	1.20/0.98	-0.41/-3.89	2.73/-2.88
0-800	0.35/0.50	1.06/0.94	-0.59/-3.76	-0.77/-2.04
0-900	0.17/0.49	1.04/0.91	-0.74/-3.74	-3.91/-2.25
0-1000	0.13/0.48	0.98/0.89	-1.08/-3.74	-7.48/-2.57
100-500	0.34/0.44	1.23/1.00	0.54/-3.40	14.21/-0.84
100-1000	**0.21/0.48**	**1.07/0.96**	**-0.39/-3.34**	

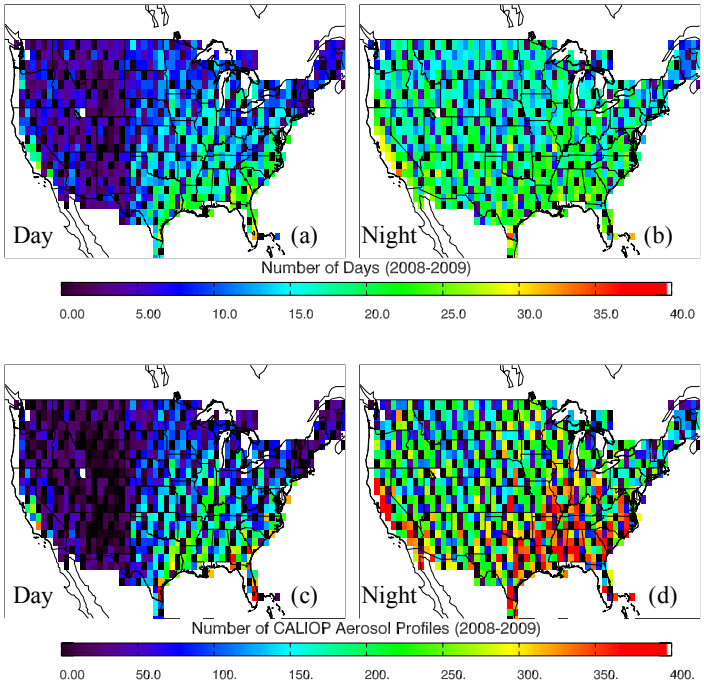
Table 2. Statistical summary of a sensitivity analysis varying the height of the surface layer, including R^2, slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in μg m^{-3}, and percent error change in derived $PM_{2.5}$, defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.

	Analysis (Day/Night)		
$PM_{2.5}/PM_{10}$ Ratio	**Deming Slope**	**Mean Bias (CALIOP - EPA; µg m^{-3})**	**% Error Change**
Low ratio (-1 STDEV) = 0.24	0.43/0.38	-7.81/-8.61	-60.00%/-60.00%
High ratio (+1 STDEV) = 0.88	1.57/1.41	5.39/0.77	46.67%/46.67%
0.6	**1.07/0.96**	**-0.39/-3.34**	

Table 3. Statistical summary of a sensitivity analysis varying the $PM_{2.5}$ to PM_{10} ratio used, including slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in µg m^{-3}, and percent error change in derived $PM_{2.5}$, defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.

Analysis (Day/Night)							
Aerosol Type				**R^2**	**Deming Slope**	**Mean Bias (CALIOP - EPA; µg m^{-3})**	**% Error Change**
	a_{scat}	a_{abs}	Γ				
Smoke	5.26	0.26	0.18	0.10/0.44	0.86/0.78	-1.81/-4.26	-11.53/-10.54
Sea salt	1.42	0.01	0.46	0.18/0.48	2.92/2.64	22.42/12.93	184.12/184.99
Dust	0.52	0.08	0.00	0.05/0.39	9.01/8.18	102.04/70.82	826.94/843.33
Sulfate	**3.4**	**0.37**	**0.63**	**0.21/0.48**	**1.07/0.96**	**-0.39/-3.34**	

Table 4. Statistical summary of a sensitivity analysis varying the aerosol type assumed in the derivation of $PM_{2.5}$, including R^2, slope from Deming regression, mean bias (CALIOP - EPA) of $PM_{2.5}$ in µg m^{-3}, and percent error change in derived $PM_{2.5}$ defined as: ((mean new $PM_{2.5}$ – mean original $PM_{2.5}$)/mean original $PM_{2.5}$)*100. The row in bold represents the results shown in the remainder of the paper.



Appendix Figure 1. For 2008-2009 over the CONUS, for each 1° x 1° grid box, the number of days and CALIOP Level 2 5 km aerosol profiles used in the creation of the maps in Fig. 3 for (a, c) daytime and (b, d) nighttime measurements. Values greater than 400 profiles for (c, d) are colored red.