

Interactive comment on “Use of the CALIOP vertical feature mask for evaluating global aerosol models” by E. P. Nowottnick et al.

E. P. Nowottnick et al.

edward.p.nowottnick@nasa.gov

Received and published: 1 July 2015

Anonymous Referee #1 Received and published: 19 February 2015 General comments This paper is good but it has the potential to be excellent. This really is a valuable study. Verifying models in comparison with observation, particularly vertically resolved observations, is vitally important for reducing uncertainty in model predictions, and yet the field is still wide open in terms of doing vertically resolved comparisons. The experiment with the “Level 2” constructed VFM applied to the model data is a very useful study and provides unique diagnostics not otherwise available. I also endorse the authors’ recognition that observations are complicated by the need for retrievals and therefore must also be validated, and therefore the attempt to make a second compari-

C1801

son using the “Level 3” constructed VFM is admirable. However, the conclusions drawn from this second comparison are relatively weak. In terms of dust and polluted dust, these conclusions seem to be entirely dependent on ad hoc thresholds and there’s no attempt made to justify why these particular thresholds should contain information relevant to CALIOP extinction estimates. The Level 3 VFM experiment does appear to have led the authors to find a legitimate algorithmic error relating to smoke and marine aerosol in the Gulf of Guinea, but this conclusion is given very little discussion or analysis. I have a serious concern about the apparent lack of awareness of any related or similar work. Although the manuscript includes a lengthy bibliography, there are surprisingly few references to studies involving CALIOP. The authors claim that there has been no prior validation of the CALIPSO vertical feature mask, yet there are at least three CALIOP validation papers focused specifically on the aerosol classification/lidar ratio selection (“Aerosol Classification from Airborne HSRL and Comparisons with the CALIPSO Vertical Feature Mask” (Burton et al.); “Evaluating CALIPSO’s 532 nm lidar ratio selection algorithm using AERONET sun photometers in Brazil” (Lopes et al.); and “Comparison of CALIOP Level 2 aerosol subtypes to aerosol types derived from AERONET inversion data” (Mielonen et al.)). Additionally, there are many papers that focus on the CALIOP extinction or aerosol optical depth but which are therefore also relevant to the lidar ratio selection. Many of these have conclusions very relevant to the current manuscript, including notably several papers like this one that comment primarily on dust and polluted dust, including but not limited to Campbell et al. 2012, Oo and Holz 2011, Schuster et al. 2013, Tesche et al. 2013, and Rogers et al. 2014. Fortunately, it is very easy to obtain a bibliography of CALIPSO related papers, including validation papers specifically, by accessing the CALIPSO web site (<http://www-calipso.larc.nasa.gov/resources/bibliographies.php>). I urge the authors to read these works and familiarize themselves with what is already known about the performance of the CALIPSO vertical feature mask in order to improve the background and deepen the discussion in this manuscript, and also specifically to refine the description of how this work is novel. The work in this manuscript is valuable and original, but it should be

C1802

put into the context of what has already been done. This lack is made more glaring by some of the authors' statements that aspects of this study have never been done before, which in some cases is not actually true. My other primary concern is an apparent confusion about the distinction between backscatter and attenuated backscatter. Both quantities are relevant to this discussion, but the manuscript doesn't accurately distinguish between them. Hopefully this is only a failure of terminology, but some of the calculations reported here depend on correctly converting between the two quantities. I hope that the authors will check their calculations if necessary, and correct the terminology and notation throughout the manuscript. I have noted some of the specific locations below, but not necessarily all of them. I am also concerned with the description of the calculation of attenuated backscatter using the model data. This definition does not appear to agree with the CALIPSO products it's being compared with. My last general comment is about quantitative comparisons. The comparisons in this paper are essentially qualitative. A more quantitative comparison could provide significant additional benefit.

We thank reviewer #1 for taking the time to review our manuscript. By collectively addressing the specific comments below, we have addressed the reviewer's primary concerns: 1. We have provided a more detailed description of our evaluation of the CALIOP VFM algorithm using our "MERRAero-Level 3" now named "MERRAero-Extinction" experiment. 2. We have included many of the citations suggested above, and have modified the text to place our study in its proper context with respect to those studies. 3. Backscatter vs. attenuated backscatter. We have carefully checked our use of backscatter versus attenuated backscatter in our manuscript and have adopted notation used in the CALIOP ATBD.

Specific comments P 1402, line 5. Why focus specifically on dust? I can find no explanation for why the methodology should be limited. If the methodology or results are only appropriate for dust, would you explain why? One of the results in your Conclusions section is not related to dust and appears to be left hanging to some extent. Have

C1803

you considered applying this study to all types more systematically?

A practical reason for this is that our model does well with dust, and dust is relatively easy to distinguish from among the various fine-mode aerosol types. Also to the point is that previous studies have identified the dust flags in the VFM as being the most reliable of the type definitions. For example, Mielonen (2009) found best agreement for desert dust and polluted dust when comparing AERONET derived and CALIOP aerosol types. Additionally, we cite Burton et al. (2013), who found that the CALIOP desert dust aerosol type had the best agreement (80%) when compared to High Spectral Resolution Lidar (HSRL) typing. There was significantly less agreement (as low as 13%) for other aerosol types, so we focused on a dusty region for this study. We have added text that cites these studies as justification for our focused analysis in the Introduction: Mielonen et al. (2009) classified aerosol types from ground-based AERONET single scattering albedo (SSA) and Angstrom exponent (AE) retrievals, finding best consistency between the AERONET-determined types and near-coincident CALIOP VFM reported types for the CALIOP dust (AERONET coarse/absorbing) and polluted dust (mixed-coarse-and-fine/absorbing) classifications. Recently, Burton et al. (2013) evaluated CALIOP VFM aerosol typing using airborne High Spectral Resolution Lidar (HSRL) under-flights, finding classification agreement in 80% of the CALIOP-identified dust cases, but generally worse agreement in the typing for other species. We also note the misidentification of marine layers as dust in the CALIOP VFM our conclusions as a bias in the CALIOP VFM algorithm, which will affect the frequency of "observed" dust layers near the surface. This bias has been published Amiridis (2013) and we have made a reference to this study.

P 1402, line 20-24. This statement in the abstract could be reworded and clarified. It is not immediately clear why a "greater presence of dusty vs. marine aerosol" indicates problems with the CALIPSO classification of dust and polluted dust.

We have clarified this sentence: "Additionally, we find that the CALIOP VFM algorithm is challenged when classifying aerosol features that are depolarizing but the contribu-

C1804

tion of dust to the total extinction is low, as our application of the CALIOP VFM algorithm to MERRAero aerosol distributions identified aerosol layers dominated by marine aerosol as dust dominated.”

P 1405, line 13. Should be “than attenuated backscatter”. An accurate (unattenuated) backscatter value, like extinction, also requires the lidar ratio. Without lidar ratio, all that’s available to compare to models is attenuated backscatter.

We have corrected this sentence to read “than attenuated backscatter”.

P 1405, line 20. “Never been comprehensively evaluated.” See general comments above. Many papers have evaluated the VFM either directly or indirectly through extinction comparisons. These should be acknowledged here, even if you do not consider them “comprehensive.”

We agree that “comprehensively evaluated” is not correct and have omitted this from the manuscript. We highlight evaluations of the VFM by Tesche et al., (2013), Mielonen et al. (2009), and Burton et al.,(2013) as a validation of the VFM and clarify that as far as we know, our study is the first of its kind because it uses the VFM itself to evaluate a model on a monthly timescale.

P 1408, line 22. Please be careful of notation. Here you use the symbol beta to indicate attenuated backscatter coefficient, but in Eq. 1 it indicates backscatter coefficient (i.e. “unattenuated” backscatter).

We have deleted “attenuated”.

P 1408, line 23. The cloud-aerosol discrimination algorithm now also uses integrated volume depolarization ratio, as of version 3.01. See the data quality summaries on the online CALIPSO user’s guide for updates to the original published algorithms.

Thank you for pointing this out. The text now reads: Once a feature has been identified, the CALIOP feature finding algorithm utilizes a five-dimensional probability density function (PDF) derived from observed cloud and aerosol properties that input the mag-

C1805

nitude of attenuated backscatter, volume depolarization ratio, latitude, altitude, and the color ratio $\chi = \beta'_{1064} / \beta'_{532}$, defined as the ratio of the attenuated backscatter at 1064 nm (β'_{1064}) and 532 nm (β'_{532}) to distinguish clouds from aerosols (Liu et al., 2010).

P 1409, line 15. Delete “attenuated”. Beta is the (unattenuated) backscatter in this equation (the attenuation is reflected in the transmittance term). Please check throughout the manuscript to make sure backscatter and attenuated backscatter are used correctly. It seems that the symbol “beta” and the phrase “attenuated backscatter” are used throughout to label both quantities, backscatter and attenuated backscatter.

Attenuated has been deleted. Also, we have checked to make sure backscatter and attenuated backscatter are used correctly. Beta is used to denote backscatter while beta prime is used to denote attenuated backscatter.

P 1409, Eq 2. In this case, the CALIOP algorithm does indeed use attenuated backscatter in this calculation according to Omar et al. 2009, but therefore a different symbol is required. Beta-prime is usual.

Beta prime is now used in equation 2.

P 1411, line 15. Cloud-Aerosol Detection Failure and Aerosol Type Failure. Please clarify. Are these set using the QA and CAD flags, or is this meant in the context of the experiment to apply a VFM algorithm to the model data? Somewhere, please say what CALIOP data versions you use, and discuss what – if any – QA flags you apply.

In Section 2.1, we note that we are using Version 3.01 data. We map to the CAD Failure and Aerosol Type Failure flags when the QA feature type and feature sub-type flags are equal to 0 (no confidence), respectively. In the text: “Additionally, we include flags for clouds and signal attenuation (no signal), the Earth’s surface, and instances when the quality assurance flags are equal to zero for the feature type (Cloud – Aerosol Detection Failure) or the aerosol feature sub – type (Aerosol Typing Failure).”

P 1413, Line 3. What are the inputs to the look-up table? Is it a single value for mass-

C1806

to-extinction conversion and one for mass-to-backscatter conversion for each aerosol species? If so, perhaps consider including them in a table. If it's more complicated, it would be nice to read a description of what other factors are involved.

The input to the look-up table is the simulated mass mixing ratio. This is converted to extinction/backscatter using mass extinction/backscatter efficiencies that are dependent on particle size, wavelength, and relative humidity. We have modified the text: "Aerosol optical properties, whether from Mie theory or the non-spherical dust optical properties database, are presented in look-up tables that provide quantities such as the mass extinction, scattering, and backscattering efficiencies, particulate depolarization ratio, and phase function, as a function of wavelength, relative humidity, and dry particle size. We straightforwardly convert our simulated mass mixing ratios to aerosol optical properties using these tables."

P 1413-1414, Section 3.2. There is a lot of jargon in this paragraph. Could this description be made more informative for those who are not already familiar with this model? For example...

Line 6. Provide a definition and/or reference for "analysis splitting". Is this term meant to encapsulate the following description of the 2-dimensional and then 3-dimensional analysis? If so, maybe just insert the transitional phrase "That is" to guide the reader.

Analysis splitting is a 2-D analysis followed by a mapping of the 2D AOT increments into 3D increments of aerosol concentration for each species. This is explained in the text, but we have modified it so that it is clearer: "The AOT analysis in GEOS-5 is performed by means of analysis splitting where first a 2-dimensional analysis of AOT is performed using error covariances derived from innovation data. Then, the 3-dimensional analysis increments of aerosol mass concentration are computed using an ensemble formulation for the background error covariance."

Line 7. Provide definition and/or reference for "innovation data". Observations minus climatology? Hollingsworth and Lonnberg?

C1807

Innovation data is the difference between the observations and the model first guess interpolated to observation locations. We have added a reference (Daley, 1991) for this terminology and have provided an explanation in the text.

Line 8. "increments of aerosol mass concentration". Does this mean an incremental correction to the first-stage analysis?

Increments used here is the difference between the analysis and the model first guess. We have added this to the text.

Line 10. Provide reference for Local Displacement Ensemble methodology.

Unfortunately, there is none for Local Displacement Ensembles. However, we refer to Daley et al., 1991 for a description of ensemble perturbations.

Line 10. "ensemble perturbations". Does this mean differences compared to the background?

Ensemble perturbations are the difference between the ensemble member and the ensemble mean. We reference Daley, 1991 again here and provide a description of ensemble perturbations.

Line 15. What are "analysis increments"?

Analysis increments mean the analysis minus the first guess. We note this in the text.

P 1416, line 12. AE reflects both particle size and coarse mode fraction. Schuster et al. 2006 "Angstrom exponents and bimodal size distribution" show how they affect it, and specifically how different wavelength pairs are differentially sensitive to coarse mode fraction vs. particle size in the fine mode.

We have modified the text and cite Schuster et al., 2006: "AE is a measure of the dependence of AOT on wavelength, which is a function of particle size (Eck et al., 1999) and fine mode fraction (Schuster et al., 2006)."

C1808

P 1416, line 24. These explanations for errors in AE could also be relevant to the error in calculated depolarization discussed on page 1413.

We certainly have some errors in the simulated particle size distribution, and this is a fundamental limitation in a model like GOCART, which does not resolve size for all species. Nevertheless, we did perform some offline sensitivity experiments to determine what contribution, if any, uncertainty in the dust particle size distribution played in calculated depolarization ratio. We found very little sensitivity, with values staying very close to 0.25. This leads us to believe that our lack of ability to simulate the observed depolarization ratio is related to some more fundamental aspect of how we do the dust optical properties. Either ellipsoidal (or spheroidal) shapes are inadequate, or else possibly it is the assumption that the particles are homogeneous by composition that leads to this effect. We explain this in the text in Section 3.1: “Sensitivity analyses we performed (not shown here) suggest that our simulated dust depolarization ratio is not strongly sensitive to any errors we have in the simulated dust particle size, as suggested by Freudenthaler et al., 2009, so we speculate here that we are fundamentally limited either because of our assumption that dust particles are ellipsoidal or because of our assumption that they are homogeneous in composition. Imaging of dust particles shows that neither assumption is true (e.g., Buseck and Posfai, 1999), and there is at least theoretical evidence that inclusion of heterogeneity within particles can lead to higher depolarization ratios (Mishchenko et al. 2014).”

P 1417, line 14. Figure 3 does not really “illustrate the impact of CALIOP aerosol typing”. Rather, it shows a disagreement between CALIOP and the model, for which one reasonable explanation is an error in typing. This is also at P 1430 Line 12 in the Conclusions section, where it says “we demonstrated” but actually it was merely stated, not demonstrated. To literally demonstrate the impact of typing would require an experiment where the typing was actually changed in the CALIOP retrieval, resulting in better agreement with “truth”. A good example of this is in a recent CALIOP Level 2 validation paper, Rogers et al., 2014.

C1809

We have reworded this sentence: “In Figure 3 we compare monthly mean MERRAero dust vertical extinction profiles to those determined by CALIOP using the VFM.” In our conclusions we have replaced “demonstrated” with “discussed”.

P 1417, line 16-19. I’m a little confused by this methodology. The VFM is a Level 2 product, not present in Level 3, but you talk about using Level 3 gridded extinctions. Can you add more explanation about how the VFM is applied to the gridded extinctions, if that’s how it was done? On the other hand, perhaps you are using the calculated dust extinction in the Level 3 product. In that case, I guess all that’s missing is a citation for the paper describing the Level 3 dust extinction product, to make it more obvious where this comes from.

Winker et al., 2013 describes this methodology, so we have added a reference to this paper. We are simply using the Level 3 dust extinction product as described in line 19: “In Figure 3, CALIOP dust extinction data are from their monthly Level 3 gridded (5° longitude $\times$ 2° latitude) product. . .”

How does “polluted dust” affect this comparison? If there is aerosol in this scene being typed as “polluted dust” that presumably means that there is additional “dust extinction” due to the dust component of the polluted dust mixture. However, it’s not clear if this is being accounted for, and if so, how.

Only desert dust aerosol layers are used to construct the CALIOP Level 3 dust extinction product. Unfortunately, the fraction of dust in “polluted dust” layers is not provided in the CALIOP data, so this contribution is excluded. In an effort to fairly compare the CALIOP Level 3 dust extinction product to MERRAero, we sample the model only where desert dust features are detected by CALIOP as explained in the text.

Section 3.3, general comment. It could be helpful to show some more quantitative comparisons. There seems to be some subtlety involved in determining where the comparison is described as good and where disagreements are highlighted. For example, you indicate good agreement with MISR right off the coast but lack of agreement

C1810

in the Caribbean. From the figures alone, it looks like it could be a similar magnitude of error off the coast compared to the Caribbean. Similarly, you say the CALIOP dust plume agrees in latitude range but not in altitude range at 7.5 degrees longitude, but the figure could be interpreted to show similar disagreement between 25-30 degrees latitude as there is between 4 and 4.5 km altitude. It's hard to read colors off the charts so I could easily be misinterpreting the magnitude of these differences. In any case, I take the authors' point that these are not serious model errors, but quantifying them systematically would be helpful.

In an effort to provide a more quantitative evaluation of MERRAero AOT, we include a new figure that compares all hourly AERONET AOT observations to MERRAero for our region of interest (20S-45N, 100W-60E) for July 2009. From 285,696 observations, we find an r^2 of 0.716 and a MERRAero bias of -0.016. Similarly, we include similar statistics for Angstrom Exponent and find an r^2 of 0.653 and a MERRAero bias of +0.115. Additionally, we refer to Buchard et al., 2015, who provides a comprehensive analysis on MERRAero biases.

Section 4. This strategy is good: using two different methods of calculating a VFM-like product from the model allows for more detailed investigation of the discrepancies. The caveat about the influence of the subjective on the "Level 3" method is important, though. You might consider pointing out here not only the difficulty in associating aerosol species to categories, but also that the selection of thresholds is necessarily somewhat arbitrary.

This is a good point. We have modified the text: "Some of the MERRAero to CALIOP VFM mapping is straightforward, such as desert dust, but in other cases—specifically those related to polluted aerosol types—the typing criteria are more subjective due to our use of thresholds that determine the presence of individual aerosol types."

Section 4.1. Is there any concern that sampling using the CALIPSO VFM could cause a bias in the comparison in cases where MERRAero may have the aerosol extinction

C1811

spread over a larger area (either horizontally or vertically) or have minor transport errors, that would lead to the VFM mask sampling only a subset of the modeled aerosol layer?

We explored different sampling strategies (sampling the model at every CALIOP aerosol feature vs. averaging to the model grid first) and found little sensitivity, as the GEOS-5 grid is significantly coarser in the horizontal and vertical when compared to CALIOP. Additionally, finding transport errors is one of the goals for constructing our MERRAero-CALIOP VFM for comparison to CALIOP VFM. By sampling MERRAero across CALIOP features and applying the CALIOP VFM algorithm, we found that MERRAero transported too much dust downwind of the Sahara.

P 1419, line 20. Combining aerosol type this way will probably be different in general than averaging the aerosol properties first and then applying the VFM logic to the averaged properties. The second method is more like what the CALIOP algorithm does, although the size of the grid box is very different. Have you considered trying it this way, to see if the result is sensitive to the aggregation method?

We explored this method as well and due to the coarse resolution of the GEOS-5 grid compared to CALIOP we did not see much sensitivity to the aggregation method, particularly in our region of interest. We have added a brief discussion on this topic in the text: "As a sensitivity test to our sampling methodologies, we explored the impact of first averaging CALIOP features to the GEOS-5 grid before applying our MERRAero VFM algorithms and found little sensitivity of aerosol typing to aggregation methodology due to the coarse resolution of GEOS-5 when compared to CALIOP."

P 1419, line 22. How many CALIOP-identified "layers" are typically combined in a single model grid box?

Typically, the number of unique CALIOP-identified aerosol layers combined in a single model grid box is on the order of 10.

C1812

P 1420, line 7. I think this calculation of attenuated backscatter is problematic. I believe the CALIOP attenuated backscatter product (at least the layer product – is that what you are using?) includes only particulate scattering attenuated by particulate transmittance. I think the molecular terms have been corrected out. Have you considered contacting a member of the CALIOP processing team to help check your definitions?

Yes, I have contacted a CALIOP team member regarding this issue. Because the CALIOP VFM is used in near real time, the total attenuated backscatter and the volume depolarization ratio (particulate + molecular) are used for aerosol typing. This is also why the estimated depolarization ratio is used, rather than the particulate. We utilize the layer product for identifying aerosol layers for sampling MERRAero.

Figure 4. The quantitative comparison of attenuated backscatter may be thrown off by the possible error above, but the comparison for this case is impressive. I am pleasantly surprised in particular at the good agreement in location of the layers in latitude, and the near agreement in layer heights. It looks like the layer top height at C is a bit too high compared to CALIOP and the marine layer between A B is a bit too shallow, but in general there is clear correspondence. The ability to validate the layer heights in the model is to my mind one of the key advances that this comparison conveys over what has been previously done with MISR or MODIS. This could be highlighted more. It would be nice to see some discussion specifically of the layer heights and depths and a vertically resolved profile comparison or more simply a comparison of the partitioning of extinction to height (perhaps as done by Koffi et al. 2012).

While a thorough evaluation of the layer heights and depths could be a separate study of its own, Buchard et al. (2015) presents a nice seasonal averages of CALIOP and MERRAero total attenuated backscatter profiles for several aerosol regimes (Saharan dust, smoke from central Africa, and anthropogenic pollution over Asia) and found good agreement between the observations and model in most regions. Additionally, given that CALIOP extinction is a function of aerosol typing, errors in the aerosol typing can have large implications for the CALIOP extinction profile (as we presented in Figure 3).

C1813

P 1421, line 15ff. Discusses a detection threshold for extinction and then converts to backscatter using a conservative marine lidar ratio. However, CALIOP's native measurement is attenuated backscatter, so it seems like it would be more direct and accurate to implement the threshold on attenuated backscatter instead, if you can find the appropriate value to use in one of the CALIOP papers (or ask the CALIOP team).

We use MERRAero attenuated backscatter to construct our MERRAero-CALIOP VFM (previously MERRAero Level 2 VFM). Aerosol optical thickness is most frequently validated optical quantity in MERRAero and is assimilated into the model, therefore, it was a logical choice when constructing a VFM that is representative of what is actually being simulated in the model. Additionally, our minimal extinction threshold is based on applying the marine lidar ratio to the minimal CALIOP backscatter near the surface and is the most conservative estimate of observable CALIOP. Therefore, any extinction layer that meets our minimal extinction threshold would meet the minimal detectable backscatter for CALIOP.

P 1423, line 8. Very true, the need for thresholds for determining how to combine model species into mixture types adds another complication. I also agree that it is nevertheless useful. So, a sensitivity study to assess how sensitive the results are to the choice of thresholds would be extremely valuable. It would help a lot to lend confidence to the interpretation of the results, if it could be shown that they are not wholly dependent on a precise choice of thresholds. This sensitivity test is mentioned as a possibility in the conclusion (P 1432, line 24), but I think it should actually be attempted, because the conclusions apparently depend so heavily on the thresholds.

This is a nice suggestion. We explored the sensitivity of our July 2009 Level 3 VFM when we varied the minimum contribution from each aerosol type in increments of 5% from and found little sensitivity above 15% (we have now noted this in the text). As a further sensitivity, we lowered the minimum threshold to 10% and found that the Saharan dust plume became broader and had a greater contribution from polluted dust. We discuss this result further in our Conclusions, as it adds strength to our findings on

C1814

the challenges of classifying aerosol layers when the dust loading is low.

P 1423, line 24. Are the values in Table 3 from mixture types from the MERRAero mapping of the whole globe or just this transect and from a day, a month, or a longer period of time?

These values are from July 2009. We have noted this in the text.

P 1424, line 26ff. The problem of modeling dust where there should be marine almost looks like it might be a problem with the data aggregation (complicated by a more minor problem with the layer height). The figures look to have a rather shallow layer of enhanced extinction and very low depolarization which might in fact be marine, topped a deep layer of very little extinction and enhanced depolarization which is probably too tenuous to be of much real consequence. Is the extinction mask applied before or after the vertical aggregation of data?

We aggregate our simulated total attenuated backscatter, depolarization ratio, and extinction using the aerosol layer top and bottom provided by CALIOP. In Figure 5, both our MERRAero-CALIOP and MERRAero-Extinction VFMs flag a shallow marine layer with a desert dust/polluted dust layer above and are consistent with the differences seen in the simulated extinction/depolarization ratio.

P 1425, line 11. This conclusion “CALIOP VFM algorithm has difficulty properly identifying aerosol type” depends very strongly on the comparison using equivalent thresholds. The CALIOP thresholds may indeed be too permissive for dust, or it may be that the model thresholds are too conservative. Both are somewhat arbitrary; the fact that they don’t agree is hard to interpret, so this statement seems overly harsh.

This is a good point. The model thresholds in our MERRAero-Extinction VFM algorithm could also be too conservative. We have modified the text to reflect this: “This highlights the challenge of assigning an aerosol type for aerosol mixtures, particularly when the quantities used to assign type are faint. Here, there was sufficiently enough dust

C1815

in MERRAero to classify the aerosol layer as polluted dust in terms of the estimated particulate depolarization ratio in the MERRAero-CALIOP VFM, while comparison to our MERRAero-Extinction VFM reveals a lesser presence of dust and a greater contribution polluted continental (smoke + sulfate) aerosols. In this case, the CALIOP VFM algorithm threshold may be too permissive for classifying aerosol features with low estimated particulate depolarization ratios (i.e. polluted dust) or it is possible that the model thresholds in our MERRAero-Extinction VFM are too conservative.”

P 1426, line 10. Given the problem with data gaps described here, if there are any remaining gaps, it would be useful to have a separate color indicating “no data” which should be kept distinct from “clear of aerosol and clouds”.

In our monthly VFM figures, we only consider aerosol features from non-cloudy profiles, so we have changed “clear of aerosol and clouds” to “no data”. We have corrected this on the figure legend.

P 1426, line 25. Looking at the online browse images for CALIOP for July 2009 (v3.01) I cannot find a marine feature like this. Could it be due to sampling/regridding? Granted, I have not spent as much time looking into it as the authors so I may have missed it. However, I think it’s unfortunate to simply say “we suspect an error” in a published dataset without providing more supporting evidence.

We have found the file containing this error and have removed it from the analysis.

P 1428, line 20. “CALIOP VFM algorithm potentially flags aerosol layers as dusty when the actual dust aerosol loading is small”. Again, what this means is that there is a mismatch between the depolarization threshold and the definition in use in this manuscript for “small”. It does not necessarily mean the depolarization threshold is insufficient for determining dust or set at the wrong level (although it may be). What is the right threshold value should be determined by what value will produce the best extinction product, since that is the purpose of the VFM. Perhaps there should be some discussion of that question. It’s not clear to me that this experiment is assessing that

C1816

question.

This point is valid. The primary goal of the CALIOP VFM is to assign an appropriate lidar ratio for converting the backscatter to extinction. However, when using the VFM type to evaluate specific aerosol types in a model as we have, the type is significant. At the reviewer's request, we tested the sensitivity of our MERRAero-Extinction VFM to our choice of thresholds and found little sensitivity in the identification of the Saharan dust plume when the individual minimal contribution threshold to the total extinction was varied from 15% to 35%. When the minimal contribution required for an aerosol species was lowered to 10%, we found best agreement between the MERRAero-CALIOP and MERRAero VFMs and serves and highlights that using estimated depolarization ratio alone presents a challenge for identifying aerosol type when there is a small contribution from dust.

P 1432, line 3-5. This seems like a more serious error and therefore a more valuable finding, since it is probably not very dependent on the thresholds used for defining model type. It deserves more discussion and really should not be brought up for the first time in the "Conclusions" section. Please note that there is already some discussion in published literature about errors due to the "elevated layer" rules in the VFM, which could be part of the reason for this error. It would be good to discuss the finding in context of what is already known about this problem. Ford and Heald 2012, Kanitz et al. 2014, Campbell et al. 2012 are some possible examples; there are probably others.

This discussion is meant to serve as a clear example of how assigning different aerosol types in the VFM can have significant implications on the extinction. We have added appropriate references (Campbell et al., 2012; Ford and Heald, 2012; Kanitz et al., 2014) to the papers suggested who have also made this point in the literature.

P 1432, lines 9-14. It's a good point that a careful experiment of this type, applying the CALIOP VFM to a truth dataset, is a very valuable exercise. There are difficulties, since it is very hard to determine what "truth" is in terms of what the VFM is attempting

C1817

to deliver, which is aerosol types relevant to an extinction retrieval, but it's still good to attempt it. However, it's not true that this hasn't been done before. See Burton et al. 2013 for an experiment applying the CALIOP algorithm including VFM to airborne HSRL data. See Ford and Heald 2012 for an experiment involving the VFM with model data used as truth to assess a related question of how aggregating in a single-type layer affects CALIOP results.

In this paragraph we are referring using the VFM as an Observing System Experiment to evaluate the CALIOP algorithm itself. While the suggested references also provide evaluations of the VFM, the algorithm itself was not directly applied as we have in our study. We have, however, made reference to these previous studies when describing previous studies that use the CALIOP VFM in our introduction and refer to others in the Conclusion.

Technical comments

P 1404, line 10. Needs rewording. The need to characterize uncertainties is universal; it's not a "limitation" of global aerosol transport models. Perhaps you mean the relative lack of such characterization to date limits the utility of these models.

Yes, we have modified the text: "However, the current lack of uncertainty and error characterization limits the utility of global aerosol transport models."

P 1404, line 15. "do not" should be "does not"

Changed to "does not".

P 1410, line 19. "tying" = "typing"

Changed to "typing".

P 1410, line 27. "integrated total attenuated backscatter is used to set the minimum backscatter threshold". Do you mean that the integrated total attenuated backscatter is compared to the threshold?

C1818

We have modified the text: “The integrated total attenuated backscatter is used for identifying aerosol type over different land surface regimes in the CALIOP VFM algorithm.”

P 1414, Line 11. “are meant to represent”. This wording is a little odd. Meant by whom? Perhaps just “perturbations represent”.

We have reworded this sentence: “This calculation is performed using the Local Displacement Ensemble (LDE) methodology under the assumption that ensemble perturbations (Daley, 1991) are used to represent errors in the placement of the aerosol plumes. “

P 1415, line 10. At what wavelength?

Added 550 nm.

P 1415, line 29. Should “except over oceans” be “except over land”?

Changed to “except over land”.

P 1416, line 15. At what wavelengths?

Added 550 nm for AOT and 440/870 nm for AE.

P 1417, line 18-20. “CALIOP extinction” vs. “MERRAero dust extinction”. I think it would be better to say “CALIOP dust extinction” since both quantities in the figure are meant to be the dust component only.

Changed to “CALIOP dust extinction”.

Figure 4. Please consider using the full colorbar range for each plot. It’s disappointing to limit the amount of information that can be conveyed by using a color bar with only 7 increments where 5 of them are essentially yellow (as in d and f). It’s difficult to see distinctions even at the critical values of 0.075 and 0.2 depolarization, that indicate differences in aerosol type.

C1819

We have increased the number of contours for this figure.

P 1420, line 16. “Southwestward” instead of “southwesterly”?

Changed to “southwestward”.

P 1420, line 21. Is it averaged to 5 s or rather to 5 km?

Yes, changed to “5 km”.

P 1426, line 15. Delete extra word “in”.

Deleted “in”.

Interactive comment on Atmos. Meas. Tech. Discuss., 8, 1401, 2015.

C1820