Atmos. Meas. Tech. Discuss., 8, C147–C157, 2015
www.atmos-meas-tech-discuss.net/8/C147/2015/

Atmospheric Measurement Techniques Discussions

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

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 comparison 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,

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

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 put into the context of what has already been done. This lack is made more glaring

AMTD
8, C147–C157, 2015

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.

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 you considered applying this study to all types more systematically?

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.

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

AMTD
8, C147–C157, 2015

that's available to compare to models is 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."

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).

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.

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.

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.

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.

P 1413, Line 3. What are the inputs to the look-up table? Is it a single value for mass-to-extinction conversion and one for mass-to-backscatter conversion for each aerosol

AMTD
8, C147–C157, 2015

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.

P 14131414, 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.

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

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

Line 10. Provide reference for Local Displacement Ensemble methodology.

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

Line 15. What are "analysis increments"?

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.

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

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

AMTD
8, C147–C157, 2015

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.

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.

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.

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 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.

AMTD
8, C147–C157, 2015

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.

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 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?

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?

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

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?

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

AMTD
8, C147–C157, 2015

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).

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).

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.

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?

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

AMTD
8, C147–C157, 2015

by 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?

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.

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".

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.

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 question.

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

AMTD
8, C147–C157, 2015

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.

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 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.

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.

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

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

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

AMTD
8, C147–C157, 2015

P 1415, line 10. At what wavelength?

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

P 1416, line 15. At what wavelengths?

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.

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.

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

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

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

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