We thank the reviewers for their comments, which we have addressed in the revised manuscript as discussed below. Throughout, reviewer comments are in blue font and italic type, and our response in regular type. There have been textual changes throughout the manuscript; the most significant additions and rewrites have been to the sections dealing with uncertainty on derived concentrations (2.3), the MODIS comparison (3), vertical distribution section (4), and conclusions.

Reviewer 1

As the authors point out, the issue of satellite retrieval of aerosol volume and number concentration in the PBL is a very important one. Indeed, one can argue that the global coverage of aerosol concentration necessary to significantly improve assessments of changes in the global radiative forcing due to aerosols is predicated on the success of such retrievals – and they are not in hand. This study well illustrates this, as the authors rightly note. Nevertheless, it is quite useful to have a sort of interim progress report on the feasibility of such retrievals and in this sense the current manuscript would constitute a worthwhile publication. However, I find the presentation lacking in several regards. First, the authors must more clearly identify the value of the comparison of MODIS retrievals with the MAN data. These data are certainly not MBL aerosol concentrations and are really based on radiometers looking upwards rather than downwards (as in satellites) coupled with microphysical models that do not differ too much from the models used in the MODIS look-up tables. The authors conclude from the comparison that the MODIS average solution may be better than the best solution. Is that it? It is worth noting in this regard that such average solutions usually DO give better results in chi-squared assessments since they strongly down-weight outliers. The price paid is, of course, less actual prognostic power for any give case. I think more needs to be done here to clearly define what is being tested by such a comparison and what is not. Second, the comparison of the MAN data with actual measurements of the MBL aerosol concentration is, in my view, flawed, and should be revised. Of course, the authors are careful to point out the difficulties in such a comparison (e.g., lack of the necessary data) but I feel that the in situ data used are not appropriate (more on this in the specific comments) and suggest a better agreement (poor though it is) than is actually the case. Finally, I do not feel that the confounding influences of such non-aerosol parameters as RH have been adequately addressed. Astonishingly, RH impacts on the aerosol have not been discussed at all. Associated with this, I think that the authors could do more to suggest approaches to deal with this admittedly difficult issue. Some of these shortcomings are quite serious. Nevertheless, the aerosol remote retrieval problem is a very difficult one and I think that the study could be made acceptable for publication if the issues I raised above can be successfully addressed. More support for this opinion is given below in my specific comments.

We thank the reviewer for their extensive and thoughtful comments. The response to each point raised above is given alongside the relevant specific comment, below. We are glad that the reviewer feels the study is worthwhile—one of the main points we wish to make is that, as the reviewer notes, what we'd like to know about aerosols is currently not in hand. There remain very significant uncertainties and our intention was that by articulating those the study would serve as a useful reference for those in the future interested in improving upon the current state-of-the-art. The reviewer makes some specific comments (below) where we had either not provided sufficient detail or perhaps not articulated these uncertainties clearly enough. We feel that in our revised manuscript we have provided a more comprehensive treatment of the subject.

2.1 Page 14961, equation 1. The equation for optical depth is O.K. but, given the ensuing discussion, I think that the authors should show everything here. By everything I mean the actual equation which relates the AOD with the aerosol size distribution and thus Cv and Cn. Hence, I would use:

(equation is malformed in comment on ACPD although we can guess to what the reviewer refers)

or some equivalent from. With this formulation one can more clearly see the problem that one is trying to address with the retrieval. Note that I have added a hygroscopic growth factor to the equation, a topic that the authors do not really explicitly address at all (more on this later).

We have expanded the Introduction with an additional equation and text to show the dependence of extinction on size distribution and refractive index, along the lines of the reviewer's suggestion. We did not include the hygroscopic growth factor in this formulation, as we prefer to introduce and discuss it later on in the revised manuscript. We did not feel the reviewer's proposed formulation of f(RH) directly in this new equation was the best way to present it, as e.g. the equation then ought to contain the dependence of e.g. refractive index on RH too, so stuck with the 'classic' formulation.

2.2 Page 14965 lines 1-27 (also bottom of P. 14967 and first lines of P. 14968) The authors here point out that they are essentially doing a fit with two free parameters, Cv,f and Cv,c. This would be an excellent place to point out, or perhaps reiterate, what the authors themselves have said elsewhere, that one can get a good fit with an incorrect model. My point here is that the authors have also incorporated other information into the model, for example the index of refraction of the two postulated modes (based on AERONET retrievals) and the modal diameters (or radii) of the two modes. The actual values of such variables may well be different from the assumed choices and, if so, the fitted Cv,f and Cv,c values will have partially compensated for this, i.e., they will not be values actually present in the MBL.

We have performed a quantitative assessment in the much-expanded section on uncertainty (now section 2.3). We perturb the microphysical model parameters (size distribution parameters and refractive index) within reasonable ranges, and present the effect on per-particle and per-volume extinction. Refractive indices were discussed in our previous model definition paper (Sayer et al. 2012, Tables 9, 10, Figure 15) and generally dependence upon the value was found to be not so strong within reasonable ranges of variability. We would also like to note that the refractive index chosen in our model was not from the AERONET inversion directly, but rather our previous work included a sensitivity study to investigate the dependence on refractive index and find values which provided the best match (using several different measures) to spectral AOD over an ensemble of sites.

Certainly the values, for example, of the indices of refraction used in the model are not in agreement with numerous in situ measurements (e.g., Sierau et al, J. Geophys. Res., doi: 10.1029/2006JD007568, 2006; Shiobara et al, Atmos. Environ., 41, 4638-4652, 2007). The authors are certainly aware of this issue (P. 14967 as indicated above) but I think that they must do more than simply mention it; rather they must at least briefly (and preferably quantitatively) assess the impact on their MAN retrievals.

The Sierau and Shiobara papers the reviewer mentions do indeed show higher refractive indices than our model. However both of these studies dealt with near-coast observations, noted in those studies to be air masses influenced by the continent, and so are not representative of the pure maritime conditions which we focus on in our study (as we specifically wish to exclude non-maritime cases). Based on our literature surveys (see the Sayer et al. 2012 model development paper), open-ocean aerosols without anthropogenic/continental components are generally found to have smaller refractive indices than the Sierau and Shiobara papers. Note that refractive index was one of the parameters perturbed in the sensitivity study described in Section 3.2 of the revised manuscript.

2.3 Page 14968, lines 11-16. I do not see how the chi-squared value for the AOD fit, even assuming that it equals __, MAN, is necessarily a good indicator of the uncertainty in Cv,f and Cv,c. As per Eq. 1 (that is, my equation 1), the relationship between AOD and Cv,f and Cv,c is complex. The uncertainty in various other parameters such as index of refraction and aerosol hydration will all contribute to the uncertainty in AOD. If there are covariances present between the various parameters in Eq 1, or if some uncertainties far exceed those in Cv,f and Cv,c , then the AOD and volume uncertainties may not even be linearly proportional. Let me give just one example. There is likely a negative covariance between f(RH) and m (the complex index of

refraction), i.e., as the aerosols grow with RH, m decreases. Such covariances could easily destroy a simple proportionality between uncertainties in AOD and the aerosol modal volumes. The authors must either more clearly demonstrate that the chi-squared values are reasonable proxies for the uncertainty in Cv,f and Cv,c or excise this claim and make some more reasonable assertion, perhaps with other, more credible, assumptions.

We intended to use this point of the analysis to illustrate the uncertainty which would arise if the microphysical model were perfect, and the only uncertainty the result of the spectral AOD measurements. As we and the reviewer note, there are other sources of uncertainty. In the revised manuscript we state this point more clearly and expand the discussion of uncertainty significantly. For our analysis, we had attempted to minimise the uncertainty arising from microphysical model parameter uncertainty through the explicit subsetting of the data to only those points likely to correspond to unpolluted maritime aerosol loading conditions. We have reworded the text to emphasise that the chi-squared test just measures one component of the total uncertainty, and we now discuss explicitly these other components.

2.4 Page 14974, lines 14-21. It is here where the authors appear to make their assessment of the value of their comparison of MAN and MODIS retrievals (of course these are echoed in the conclusions). They conclude that the average MODIS solution is in better agreement with the MAN data than the "best" solution. This is something, I suppose, but not really a surprise. Unless one has strong variance in the underlying data (and restriction to marine cases severely attenuates this), an average solution will nearly invariably produce the least total residual. However, I think that the key question for any such comparison as that done here – which the authors themselves are clear to note is not a definitive evaluation against the ultimate target parameter – is what do we gain by it? Reiterating my remarks in the recommendation section, the comparison is between integral measurements looking up with fairly similar integral measurements looking down, followed by the use of, essentially, look-up table fits with rather similar parameters to retrieve the target column aerosol values. What is different between the two approaches and what does the agreement (modest, but I agree that it is there) tell us? For example, looking down, the satellite detector must deal with such issues as surface glint that the MAN photometers are not subject to. Agreement suggests that the glint issue has been successfully addressed. What else? This is the question that the author should be going all out to explicitly addressing here and I do not really see it. There are some piecemeal, partial attempts to deal with the issue at various points in the text but no coherent and definitive assessment is offered. It is needed.

We have expanded the discussion in this section. We feel that it would not be appropriate to draw too strong conclusions with respect to number concentration because of the large uncertainties in the MAN-derived data; we cannot validate the MODIS estimates, but we can provide an alternative estimate and examine consistency, which has not been done before to our knowledge. This has been emphasised again in the conclusions. This is one reason why, for example, we restrict to only the 'pure maritime' cases—we do not feel we could say much at all with confidence for the other subsets.

MODIS has provided among the most widely-used and state-of-the-art satellite aerosol datasets for more than a decade now, so we feel this comparison of Collection 5.1, which has been a community standard for maybe the past five years, and Collection 6, which is likely to fill the same role for five or so years to come, is warranted and useful on its own. We feel that one of the main conclusions from our comparison relating to number concentration (different preferential choice of fine modes between the two MODIS sensors) is an important result which will be of interest to algorithm development teams as well as some elements of the data user community.

One conclusion we have re-emphasised is that, in terms of AOD, Collection 6 appears to perform better than Collection 5 for these cases. As the clean conditions are those in which cloud contamination and surface reflectance errors are likely to cause the largest problems, this suggests that the algorithm improvements implemented in Collection 6 have been largely successful. We also wish to mention that the comparison with MODIS data is not the main focus of this study, and a full validation of the Collection 6 AOD is a matter outside of the scope of this study and would dilute the focus. So, we disagree with the reviewer that much

more needs to be done in this section, and note that the other reviews did not request these type of changes be made.

2.5 Page 14973, lines 22-25. There is a misstatement here, likely a typo but it is confusing. The authors state that the fine-mode per particle extinction is tow orders of magnitude greater than that for the coarse mode. It is of course the opposite that is true, as can be seen in Table 1.

The reviewer is correct that this is a typographical error, and should read 'smaller'. This has been corrected in the revised manuscript.

2.6 Page 14974, line23 to p. 14975, line 12. The authors commence here their comparison of MAN-derived surface aerosol concentrations with actual in situ measurements of aerosol number concentrations in the MBL. The first issue that arises in this attempt is to somehow relate the column concentrations, or burdens, with the surface concentrations. To do this, the authors adopt the assumption of exponential decreases in aerosol concentration with altitude, leading to the simple and well-known relationship between surface concentration, column burden, and scale height. For this purpose, a reasonable assortment of studies from which scale heights can be derived for marine aerosols is given in Table 7. The authors also note that one does not always have exponential aerosol profiles in the marine atmosphere. I would agree; aerosol layers aloft, for example, are quite common in the marine atmosphere (e.g., Clarke and Kapustin, J. Atmos. Sci., 59, 363-382, 2002; Kuzmanoski et al, Aerosol Sci. and Technol., 41, 231-243, 2007). Indeed, they commonly occur near one of the key sites (Lanai) used by the authors for the derivation of their aerosol model (Sharma et al, J. Atmos. Oceanic Technol., 28, 418-425, 2010). So, what does this signify? The authors do not pursue the issue, which is understandable since there is really no way to actually retrieve the surface concentrations without a simplifying assumption such as the exponential distribution. I think the authors should explicitly acknowledge this and also offer some assessment – admittedly it will be rough – of how common the exponential profiles actually are. Furthermore, there are also substantial vertical variations in other key parameters such as composition (and thus index of refraction and hygroscopicity) and, of course, RH. One compositional example would be organic aerosols, largely secondary , that lead to much of the free tropospheric aerosol mass over the oceans being organic (cf. Murphy et al, J. Geophys., 111, doi: 10.1029 /2006 JD007340, 2006; Thomberry et al, J. Geophys. Res., 111, doi: 10.1029/2009JD012721). These issues need to be explicitly acknowledged and at least briefly discussed.

We mentioned in our original manuscript that all we can reasonably do is make an assumption based on the existing measurements, to obtain results in an average sense, and comment upon the uncertainty this introduces, which was the reason for including both of those tables. We have modified the text to emphasise this point more strongly in the revised manuscript, and extended the discussion. Again, through restriction to the pure maritime subset (not just considering all data over ocean) we aim to minimise the impact of e.g. transported aerosol layers and vertical heterogeneity of composition in general, as many (although not all) examples of such elevated layers are transported (i.e. non-marine) aerosol.

We added the Sharma et al (JAOT 28, 2011) reference to Table 8 of our revised manuscript as its Figure 2 shows an example of a vertical profile fairly invariant with altitude from about 100 m to 1 km. However, we do not see the evidence that the reviewer mentions for that study showing elevated aerosol layers near Lanai. The one example of a distinct scattering layer structure we found in that work was the upper portion of Figure 4, which is stated in the text to belong to a cloud base, rather than a distinct aerosol layer.

We do not believe it is possible to make any estimate of how common exponential profiles are with any degree of fidelity—and in any case, the exponential profile will be, at best, an approximation. The field campaigns are all limited in spatial/temporal extent and so trying to quantify how comment they are from those could be misleading. The fact we found numerous examples of both near-exponential profiles as well as non-exponential profiles in the literature we surveyed indicates neither situation is uncommon. We don't believe anyone knows what the true answer is in this case, and that's precisely why we presented Tables 7 and 8 containing a range of values from the literature. We do not believe chemistry transport models can presently

address the issue, either, due to the large uncertainties and diversity remaining in them (cf. Yu et al, JGR, 2010, which we cite, and Koffi et al, JGR, 2012, which we have added a citation of in the revised manuscript). CALIOP has been used recently as an evaluation tool for models, although there remain difficulties, particularly during daytime, related to AOD/type identification. We have tried to state this more clearly in the revised manuscript.

2.7 Page 14975. There is really no specific place to cite the inadequate discussion of RH effects on the aerosol retrieval since the issue is not discussed at all. However, this issue will become most acute when one is trying to compare retrievals with actual MBL aerosol concentrations so I choose to raise it at this point in the text. The first point to remember is that RH has a marked effect on aerosol size and thus on aerosol volume but only a limited impact on particle number (essentially by moving small particles into a detectable size range). This is so well established that I hardly think I need support it. Nevertheless, let me cite at least one study (see also the numerous citations in this reference) that shows the strong impact of RH on light-scattering, namely, Carrico et al (Tellus, 52B,694-715, 2000), for various aerosols, including marine. Given this, the number-to-volume ratio will be a strong function of RH and, of course, modal radii and, to a somewhat lesser extent, the modal standard deviations of the aerosol size distribution will be impacted. This is simply physics and you cannot get around it. So how do the MODIS and MAN retrievals deal with this issue? For MODIS, it is fairly clear that the issue is simply not really dealt with (Remer et al, 2005; 2008, as cited by the authors). The MODIS LUT simply includes a number of aerosol size distributions associated with water soluble or "wet" aerosols. If these are selected in the fit, then one could conclude either that the aerosols were more hydrated than otherwise or they had differing dry sizes. Clearly there is a mixing of actual differing dry size parameters with varying degrees of hydration. For AOD retrievals one can more or less get away with this. As the authors themselves point out, RH is a function of altitude. Hygroscopicity is also a function of altitude in the marine atmosphere, and elsewhere (cf. Hegg et al, Geophys. Res. Lett., 29, doi: 10.1029/2001GL014495, 2002). And, of course, aerosol sizes change with altitude in the marine atmosphere (Clarke and Kapustin, J. Atmos. Sci., 59, 363-382, 2002). All of these things are mixed together in the AOD as per Eq. 1 (mine) so it is not too surprising, given expected covariances, that one can get at least reasonable AOD's with no explicit RH dependence – but microphysical aerosol properties recovered from such a LUT are not likely to be accurate. The MAN retrieval, as given in Sayer et al (2012, as cited by the authors), at least discusses the issue somewhat but again the authors do not seem to appreciate how difficult it is to take partial derivatives of compound variables in the real atmosphere. They examine the dependence of AERONET retrieved size distributions on RH as derived from either NCEP reanalysis (sfc, i.e., 10m) or shipboard data, in other words, surface data. These are clearly not the RH's the column aerosol sees. (Incidentally, perhaps in future work, the authors should consider the use of GCM's or CTM's to get vertically resolved RH data as per, for example, van Donkelaar et al, Environ. Health Perspect 118, 847-855, 2010.) Furthermore, the size distributions with which the RH comparison is made are AERONET retrievals and, as with MODIS, a number of variables that plausibly vary with altitude – and RH – are incorporated into the retrieved AOD's. Because the retrieved size distributions inferred from the AOD's show little dependence on RH in the 60-80% range characteristic of the lower MBL, the authors essentially neglect any RH dependence. Once again, I think that you can get away with this lack of dependence, or at least weak dependence, for AOD's but definitely not for the surface size distributions and thus Cv and Cc values derived from them – and certainly not for MBL aerosol number concentrations derived from them. I feel that a much more extensive discussion of the issue is necessary. The authors can perhaps make a case for their approach based on current necessity and the extensive averaging that they do, but it must be made plain (in my view) that this is a very problematic approach for estimating MBL aerosol number and volume concentrations.

We had neglected an explicit treatment of relative humidity (RH) in our original submission for several reasons. In our previous Sayer et al. (2012) aerosol model paper, we found ~90% of AERONET inversions were reported by NCEP reanalysis data to have boundary layer RH in the range 60%-80%. In that work we also compared NCEP with ship-borne RH, found reasonable agreement, and a similar range of values. This is mentioned in the revised manuscript. Any model dependence was small compared to the uncertainty in the RH data we had. The reviewer is correct that as we have only columnar AOD we don't know what humidity the aerosol is seeing. Additionally, because of hysteresis, it is also important what the prior history of the air mass was. We appreciate the inherent difficulties in this approach and do not see that there is a great deal to be done

about it at this point, other than do the best we can and report an honest assessment of uncertainty. When revising the manuscript, we present a more thorough and clearer treatment of RH effects, reflected in the expanded uncertainty analysis and conclusions. Since we explicitly subset to unpolluted maritime aerosol, we believe that a reasonable first-order assumption is that the boundary-layer RH can be used as a proxy for the effective RH which the aerosols are exposed to at that time (and also that the vertical extent and so heterogeneity of other factors is lessened for such aerosols as compared to e.g. dust or smoke). However this may not be true for other aerosol types (another reason to focus on the `simpler' case of pure marine aerosol).

2.8 Pages 14976-14977 (Discussion of Figure 10) The last, and most challenging, analysis offered by the authors is the comparison of the MAN derived MBL aerosol concentrations with direct measurements in the MBL. The authors derive actual aerosol concentrations in the MBL from various sources to compare with zonally averaged Man retrievals. There is a great deal of spatial and temporal averaging involved in this comparison but I do not find it unreasonable for a rough estimate of bias, though perhaps this could be more clearly stated.

We have reiterated this point (that because of sampling difficulties and lack of correlative data, it is only possible to present an illustrative comparison in an average sense) with an additional paragraph in the revised manuscript.

My main concern here, in addition to the RH issue just discussed in the previous comment (which will be most acute here), is that it is not clear to me that the authors are comparing the right quantities. First, it must be remembered that MODIS – and MAN - are retrieving first the fine mode scattering signal and, based on this, a volume fraction that is associated, essentially, with a given number-to volume ratio. Hence the interest of the authors in pointing out that these ratios (as per Table 1) are similar to those observed from in situ measurements (p. 14970, lines 7-11). However, the fine mode scattering is dominated by the accumulation mode, down only to particle diameters of perhaps 100 nm, or even larger. The smaller particles in the lower Aitken mode or nucleation mode scarcely impact the scattering signal at all in the marine atmosphere. In accord with this, the number –to-volume ratio used in the MAN model – and the observed ratios with which it is (sort-of) validated – are for the accumulation mode. Given this, when particle concentrations are derived from the MAN retrieval, they should be compared to in situ measurements of the accumulation mode. This does not appear to be the case in Figure 10. The data from the GAW network are CN data. Similarly, the data from both the Heintzenberg et al study and the Bates et al study are essentially CN concentrations (sum of all DMA channels and thus down to the detection limit of the actual detectors – which are CN counters). These counters (typically TSI 310's or 3020's) measure below 20 nm. In a way, the authors are making things difficult for themselves in trying to recover concentrations of particles that their instrument does not "see." The more immediate issue, however, is that they are looking at much higher concentrations from the in situ measurements than they should be. Both the Bates et al study and that of Heintzenberg et al report accumulation mode concentrations in addition to total number concentration and I think that the authors should do a comparison with these concentrations. In principle, I would expect to see higher bias but perhaps better correlation between the in situ and MAN values. In any case, the extant comparison is not, in my view, valid.

We agree that this is a difficult comparison to make, and that's one of the points we are trying to illustrate: these things are not well-known, and with the assumptions which must be made, we're left with large uncertainties which are difficult to resolve. It would be great to have a setup where the relevant quantities were measured simultaneously and some of these things could be nailed down more definitively, which would help out e.g. the remote sensing and modelling communities. By illustrating the current difficulties we hope this study will aid others to find the direction forward for future measurements.

We have consulted with GAW reference documents and J. Ogren (PI of two of the GAW sites used), which confirm that it is the total aerosol number concentration, rather than just Aitken mode, which is measured at these GAW sites. The Clarke and Kapustin (2002) paper we cite talks about detection limits from these types of instrument, stating the lower limits (dependent on model) are typically 5-7.5 nm radius and the upper limits

1-1.5 microns (the paper talks in terms of diameter but we talk in radius here to be consistent with the rest of our study). The 3-sigma lower tail of our fine mode in our microphysical model number distribution is ~16 nm. So, some of these particles will be counted as part of our fine mode, with the amount dependent on how close in size they are to the in situ counter detection limit.

So, we believe that comparing total number concentration is the correct thing to do here. We have run additional Mie calculations to verify, and although the per-particle extinction for typical Aitken mode aerosols is indeed much smaller than for accumulation mode, their large number (cf. Heintzenberg, Bates and other references) increases their contribution to the total AOD, although it does remain low. In a similar sense, the coarse mode is much less numerous in absolute aerosol number but this is compensated for by the much higher per-particle extinction. We are considering a bimodal lognormal volume distribution without an explicit Aitken mode, but again going back to the source AERONET inversions, a contribution from Aitken particles will be reflected in the parameters of the fine mode. So, it would be wrong to neglect the Aitken mode entirely. As the author notes, our number-to-volume ratio from the microphysical model used are in good agreement with observations. We have extended the text to mention the incomplete Aitken mode explicitly as another potential uncertainty source.

The issue of the bias between the two datasets is an important one. Due to methodological differences, the MAN estimates are inherently more likely to be high-biased due to e.g. cirrus cloud contamination, and the ground-based estimates are more likely to be low-biased due to incomplete sampling of the aerosol size distribution, which we mentioned in the original manuscript. The text has been rewritten in the revised manuscript to make this point more clearly. We found in Quant et al (J. Aerosol. Sci., 1992) a discussion of the collection efficiency of TSI models which suggests that this is less of an issue for the fine particles, and now cite. We also note that our estimates are in a similar range to model estimates in the recent study by Moore et al., ACPD 12, 2012. Admittedly the Moore study is still in review and the results are model simulations rather than observations, but their Figure 1 appears reasonably in-line with our results in terms of latitudinal dependence of number concentration.

In the revised manuscript, we now present a more general vertical profile with a well-mixed layer of arbitrary depth topped by an exponentially-decaying profile. We show that in this case only the sum of the mixed-layer height and scale height is important (e.g. surface concentration for a scale height of 1 km above a mixed layer of depth 0.5 km is the same as no mixed later with a scale height of 1.5 km, or a well-mixed layer 1.5 km deep with no aerosol above it). We estimate (again, in an average sense) what sort of total height would be necessary to bring the MAN and in-situ surface concentration estimates into a more-or-less unbiased agreement. It turns out this is about 3 km in the tropics/northern hemisphere: it is an open question whether this is realistic or not. This is, again, illustrative of the fact that these profiles are only an approximation.

Another issue worth pointing out, though it is not so critical since the authors did not choose to do the relevant comparison, at least explicitly, is that I would not expect retrieved Cv,f values to agree with those from the in situ measurements. Such a comparison might have been made, for example, to try to isolate where discrepancies are arising in the retrieval. If they had been made, it would be well to remember that both the measurements reported by Heintzenberg et al and those by Bates et al are at reduced RH (_40% or less and 55% or less respectively) whereas the retrievals are implicitly for RH's in the 60-80% range. The authors' own analysis suggests there will be decided differences in the volumes between such disparate RH's.

We chose not to present a comparison of at-surface aerosol volume in the original or revised manuscript. We felt that the number comparison was sufficient to be illustrative of the difficulties in converting columnar to surface aerosol burden (one of our main points); additionally, as the reviewer mentions, with volume there are issues related to converting between ambient and dry RH which lead to additional uncertainty. We feel that such an analysis would have increased the length of the manuscript significantly without imparting much new information to the reader. However if, in the future, situations can be found in which MAN (or other) data can be used to infer surface concentrations with some more confidence, comparison of volume as well as number may be more instructive.

Reviewer 2

In this paper "Estimating marine aerosol particle volume and number from Maritime Aerosol Network data" by Sayer et al., using MAN data, the authors estimated aerosol particle volume and number concentrations based on an algorithm they developed from a previous study for AERONET data. The aerosol physical properties derived from this study were also inter-compared with the values reported from the version 5.1 and version 6 of MODIS aerosol products. Lastly, the authors inter-compared the estimated surface aerosol concentrations with values documented from existing studies. As the authors mentioned themselves, there are potentially significant uncertainties in this study, as uncertainties from various source of inputs for the study are rather large (see below for details). However, this study has its merits, and the authors are honest about limitations of their study. Therefore, I recommend publication of this paper if the authors could address the issues below.

Our intention with this study had been to assess the current capability with regard to estimating aerosol number/volume from spectral aerosol depth, and the associated uncertainties (which can be significant). We hope this work will spur on further research. We feel that in the revised manuscript, through incorporation of the reviewers' comments, we have presented a balanced view.

The other two reviewers have made very nice comments. I will not repeat their comments. My biggest concern is the uncertainty analysis. Retrieving aerosol particle volume and number based on the spectral patterns across the selected wavelengths is not an easy task as various assumptions are needed for the study. For example, the authors used the MIE calculation which implies that clear marine aerosol particles from their studies are assumed to be spherical. However, sea salt aerosol particles may be irregularly shaped (as also pointed out by the first reviewer). Therefore, the question is how much is the uncertainty caused by the aerosol non-sphericity? Similarly, uncertainties in the study could be caused by uncertainties in other parameters such as size distribution, refractive index, relative humidity (also pointed out by one of the other reviewers), as well as spectral AOD values. Therefore, it would be a contribution to our knowledge if the authors could perform theoretical calculations and list the uncertainties in their retrievals with respect to the input parameters or major assumptions from their study. I am not trying to give the authors a hard time, but I would argue that such a table would shine a light into future research of this study area. However, if the authors think it is a daunting task, at least the authors should highlight parameters that could potentially introduce large uncertainties in this study.

The answers to these comments are contained within the responses to Reviewers 1 and 3; in brief, we have significantly expanded the uncertainty analysis (in terms of discussion and quantitative analysis), which we feel answers these points, and show that nonsphericity effects are not significant for marine aerosol.

Page 14965, line 10, it would be beneficial to readers if the authors could define Cv,f, and Cv,c here, although the two terms have been defined in Figure 1.

This has been done in the revised manuscript.

Page 14977, line 15, "artefacts" should be "artifacts"

We believe this is a difference between British/international and US English (more commonly 'artefact' in the former, 'artifact' in the latter). As the lead author is British, and Copernicus journals a European publisher, we prefer to use British English throughout.

Reviewer 3

This paper reports two sets of results, both being very instructive. Firstly, they use multi-spectral extinction measurements and a pre-determined maritime aerosol model to retrieve such important aerosol characteristics as column-integrated number concentration and average particle volume. Secondly, they compare their results with the corresponding MODIS products to provide the first accuracy assessment of the latter.
The paper is very well written as well as logically organized and should be published. However, there are some issues that the authors should at least discuss.

We are pleased that the reviewer sees the value in our work, and have expanded our discussion to reflect the points raised below.

1. As the main criterion of adequacy of their maritime model, the authors quote its ability to accurately reproduce the AERONET optical thicknesses. However, unlike extinction, the number concentration retrieval is extremely sensitive to the assumed or retrieved size distribution. This problem is well articulated in
H. González Jorge and J. A. Ogren (1996). Sensitivity of retrieved aerosol properties to assumptions in the inversion of spectral optical depths. J. Amos. Sci. 53, 3669–3683.
M. I. Mishchenko, L. D. Travis, W. B. Rossow, B. Cairns, B. E. Carlson, and Q. Han (1997a). Retrieving CCN column density from single-channel measurements of reflected sunlight over the ocean: a sensitivity study. Geophys. Res. Lett. 24, 2655–2658.

Neither paper is mentioned in this manuscript and in the preceding paper Sayer et al. (2012). It would, therefore, be very interesting to see if the quoted estimates of number concentration uncertainty derived from extinction measurements are consistent with previous findings, especially those in Jorge and Ogren (1996).

We are grateful to the reviewer for bringing these papers to our attention, which had been omitted from this and our previous study only because we were not aware of them. We have read them and mention them in the revised manuscript.

In the case of Gonzalez Jorge and Ogren (1996), their study was different from ours in that they retrieved binned size distributions assuming no special knowledge about the aerosol type. In our study we focus on the pure maritime subset, and hence are able to apply more constraints on the size distribution (which we now present associated uncertainties for in the revised manuscript) than they did. This consequently means that, although similar general conclusions are drawn, our uncertainties for this specific application are smaller than for their general case. However, their results are an important and relevant reminder that if you don't have a reasonable idea about the aerosol properties beforehand, these analyses become very uncertain. We also take care to reiterate this point in the revised manuscript (uncertainty section and conclusions).

2. The authors use throughout the Mie theory, thereby assuming spherical particles. This approach appears to be justified given the standard presumption of a relatively weak dependence of the extinction cross section on particle shape. Yet Fig. 4 of Mishchenko, M. I., L. D. Travis, R. A. Kahn, and R. A. West (1997b). Modeling phase functions for dustlike tropospheric aerosols using a shape mixture of randomly oriented polydisperse spheroids. J. Geophys. Res. 102, 16831–16847 shows a residual shape dependence of the order of 5-10% varying with particle size and size-distribution width. The authors should at least discuss whether this dependence can affect number concentration retrievals.

We believe that the spherical assumption for marine aerosol is warranted, given the fine mode is largely a mixture of hydrated sulphates, organics, and sea salt, and the coarse mode at typical relative humidities of 60%-80% as found for MAN cruises (see e.g. Sayer et al 2012, Figure 5 and Table 6) is likely to be spherical through hygroscopic growth. The only observations we are aware of for nonspherical sea salt particles is

confined to dried particles (in heated collectors) rather than ambient conditions (where humidity is significantly higher), e.g. Chamaillard et al., JQSRT, 2003, 2006. Scattering properties of sea salt aerosol in typical ambient humidities have been shown to be well-modelled as spheres (e.g. Quinby-Hunt et al, Appl Opt, 1997, section 5C). Thus, we do not believe nonsphericity is an issue for marine aerosol over the open ocean. This assumption is also made in, to our knowledge, all major satellite and chemistry transport model datasets for marine aerosols. It will, however, affect dust aerosols (which are not the focus of our study). We have mentioned this point in the expanded uncertainty discussion in the revised manuscript.

Indeed, the results of Jorge and Ogren (1996) imply that refractive-index uncertainties can play a substantial role, their effect on extinction being roughly of the same magnitude as that of nonsphericity.

Refractive index uncertainties are among those quantified in the expanded discussion which is now section 2.3 in the revised manuscript. We would also like to note that for our specific example, as we are subsetting to only that aerosol likely to be pure maritime, our uncertainty is smaller (as we have this 'a priori' knowledge of type). However in a general case if one were just taking spectral AOD and trying to determine aerosol number without any special knowledge of aerosol type, yes, the uncertainties are much larger. These aspects are discussed in the revised manuscript.

3. Unlike the extinction-based retrievals, the MODIS retrievals can be expected to be strongly affected by nonsphericity and refractive-index uncertainty. The authors may want to mention this as one of the reasons for relatively large MODIS number concentration errors.

Following this and Reviewer 1's comments, we have expanded the discussion on the comparison with MODIS data. However, as previously mentioned, we do not believe that nonsphericity effects are significant for the specific case of unpolluted maritime aerosol. Also, we would like to point out again that, since the uncertainty of the MAN-derived concentration may be significant, we cannot directly assess the error in the MODIS number concentrations through the comparison, only the level of consistency between the two datasets.

4. The authors conclude, at least implicitly, that number concentration retrievals from space are problematic. This appears to support the earlier conclusion of Mishchenko et al. (1997a), which the authors may want to acknowledge and discuss.

This paper is mentioned in the revised version of the manuscript: although its focus was on satellite data, it is relevant here in the context of our comparison with MODIS, and illustrates the point about the limited information content of current radiometers.