Andrew Sayer

I do not have any major problems with the paper, although there are a few places where I think a small expansion to the analysis/text would improve the utility of the study. As such, I favour acceptance after minor revisions. If it is determined that a second round of reviews is necessary, I am happy to review a revised version. Specific comments follow below.

We thank Dr. Andrew Sayer for his constructive suggestions.

SPECIFIC COMMENTS

P12797, line 15: I am not sure that Toth et al. (2013) is the best reference in regard to 'early studies' here. I would suggest maybe Smirnov et al. (AMT, 2011). There may be earlier references which escape me at the moment, but I don't think there was much in the way of high-latitude open-ocean AOD validation until MAN.

We thank the reviewer for these suggestions. We have added Smirnov et al. (AMT, 2011) in the reference list and made changes in discussions

P12797, line 26: Several other algorithms had been using wind speed data for a few years before these examples (although they are less prominent than MODIS or MISR), see e.g. Sayer et al. (2010) for (A)ATSR and Sayer et al. (2012) for SeaWiFS. Possibly also worth mentioning here is the recent NOAA VIIRS aerosol product, which also does incorporate wind speeds (Jackson et al. 2013).

We thank the reviewer for these comments. We have added these studies to the paper.

P12799, line 11: I think Koepke (1984), cited later in the manuscript, is a better reference than Vermote et al. (2006) here. The Koepke work serves as the basic approach used by 6S and others. On a related note, I think that the 6S paper (Vermote et al., 1997) might be better to cite than the code user guide (Vermote et al. 2006) when discussing 6S in the paper, but this is up to the authors.

We appreciate the comments from the reviewer. We have modified the text as suggested.

P12800, line 16: The rationale for looking at 0.66 μm makes sense. However why not also look at e.g. 0.87 μm? This is another key channel present and used for AOD retrieval over ocean by most sensors (MISR for example uses only 0.66 and 0.87 μm over ocean), as well as having ground-truth AOD from AERONET and MAN, and if the authors already have the required data matched up hopefully it is not too complicated to run the RT model at a second wavelength? This would help give some first order idea about the spectral dependence of the effect, which the authors note is the logical next step. I think that this would be a valuable addition to the manuscript, which should not take too much effort, and could be accomplished without excessive length (if necessary key results could be summarized in a table or something). My impression based on Figure 2 is that the effect will become negligible, but it'd be nice to have some numbers behind that, and see also my later comments on Section 2.3 and Figure 2.

Point is taken. We have updated Table 1 to also include the I and J terms for a few selected wavelengths. The 0.87 μm channel is not included in Table 1 as suggested from the reviewer, the effects are rather insignificant.

*P12801, line 20: The MODIS uncertainty should be formatted as +/-(0.03+0.05*AOD), not 0.03+/-0.05*AOD. Strictly, there is a difference between the two, and the former is correct (although the latter is often written). Additionally, it is worth noting that the team now believe that the error is larger than this (Levy et al. 2013), based on preliminary validation of Collection 6 results, which do include ancillary wind data.*

We thank the reviewer for the thoughtful comments. The MODIS uncertainty has been reformatted within the paper.

P12801, line 25 onwards: It may not make much difference for this study, but are you using Collection 5 or Collection 6 level 1 and cloud mask data? Also, this should be referred to as MYD35 rather than MOD35 as Aqua is being used.

We thank the reviewer for the thoughtful comments. Changes are made as suggested. For this study we used Collection 5 to be consistent with the work done by Toth et al. (2013). We have mentioned our rationale in the text.

On the topic of cloud masking, it would be good to state exactly how the MODIS cloud mask was used (I am not familiar with the C6 product but in previous versions at least there were both 'confidently' as well as 'probably' clear and cloudy categories). My guess is 'confidently clear' based on P12802 line 24 but it would be good to be explicit. The data version and categorization points may contribute to some of the scatter in Figure 3.

Also, we did use only 'confidently clear' pixels. We have added the discussions in the text.

I would also suggest taking an average rather than closest pixel for the MAN-MODIS-AMSR-E matchup (40 data pairs) because you have a 30 minute time window and so I think that you can go larger than taking a 500 m MODIS pixel. In fact my suggestion for this part would be to use the 0.3° lat/lon threshold and, in Figure 3, plot the median MODIS reflectance (rather than closest pixel) and put on error bars corresponding to the standard deviation or something. That will make Figure 3 a bit more informative about how much of the scatter can be attributed to RT model or calibration issues, and how much can be attributed to sampling.

We have revised Figure 3 as suggested. However, instead of a 0.3° lat/lon box, we utilized a 0.1° lat/lon box and plotted the mean radiance. Also, we added the standard deviation as suggested to represent error bars. Thanks for the suggestion.

P12802, line 10: which AMSR data product and version are used? Is this the latest version 7 (http://www.remss.com/missions/amsre)? This information should be given in the paper.
The Version 7 of AMSR-E was used. This information has been added to the paper.

P12805, section 2.3 and Figure 2: It would be good to add some context by making a statement about what fraction of the total TOA reflectance these bubble rafts contribute, i.e. show/state $\Delta R/\pi$ in terms of percent of total reflectance (for a typical scene) as well as in absolute units as given in Figure 2 presently. At the moment it's hard to get a feel for how small these small numbers are. Figure 3b doesn't do the job here because, if I understand correctly, the bubble raft contribution is folded in with the other surface contributions (e.g. glint, whitecaps). The other issue with Figure 2 is that although the absolute contribution from rafts appears to decrease as wavelength increases, Rayleigh and aerosol scattering do as well. So the relative contribution to the signal may not be decreasing as fast (and could even be increasing) at longer wavelengths.

Point is well taken. In fact, by taking the difference between the R/π term with and without including submerged oceanic bubbles, Figure 2 shows the effect of submerged bubbles on the reflectance term. We have added a new figure, as suggested, to show the relative contribution from whitecaps and from the submerged bubbles.

In the same section, is Equation 2 (with Equation 3) all that is needed (with the coefficients in Table 1) to include the bubble raft contribution at this wavelength, or is there something more? I think this should be more explicit. If this is all, and the L/J coefficients are easy to find/calculate for other wavelengths, then stating this more up front would be good to encourage people to incorporate rafts into their surface models.

In regards to the equations, Equation 2 is used to calculate the L/J coefficients and Equation 3 (which is a modified version of Equation 1 to include results from Equation 2) is the full equation needed to obtain impacts from subsurface bubbles. We have added results from other wavelengths as suggested.

Figure 3b: What is the x axis here? I think that the x axis labels should be removed. In fact I think that 3b itself can be removed, and only 3a is necessary.

We removed Figure 3b as suggested.

*Figure 4: It would be good to add some horizontal and vertical error bars to the triangles, so we can see the variability in AOD error and in wind speed within each bin. Also, the figure and caption could be clearer: are the triangles/black line the average AOD absolute error, or the average AOD difference (bias)? From the text it sounds like the error, but from the plots the error looks smaller than the cited validation results suggest (e.g. +/-0.03+0.05*AOD).*

Error bars have been added to the triangles to aid in visualizing the variability in AOD error.

Figure 5: The y axis and start of caption say MAN-MODIS, but the later sentences say MODIS-MAN. I think that MODIS-MAN is the logical way to plot things (i.e. bias of MODIS relative to MAN as a 'truth'), so suggest this is redrawn as MODIS-MAN.

We thank the reviewer for their comments. Changes have been made so that the labeling in the figure and text is consistent.

Figure 6 and related text: Is there any seasonality to the correction, driven by seasonality in wind speed? It would be good to state/show this somewhere.

We are thankful for the comments from the reviewer. Seasonal variations of wind speeds have been studied and discussion has been added to the paper.

There's one more point which is not directly addressed but may be relevant. As discussed, whitecap fraction parameterizations also depend on wind speed, and so when retrievals include wind speed data (as opposed to assuming a single value) they get better because the surface reflectance assumed gets better. However there remains a large amount of scatter in both parameterizations of whitecap fraction as a function of wind speed (I think about an order of magnitude), as well as the spectral reflectance of the whitecaps themselves (I think about 50 %). These result both from the uncertainty on the measurements but also presumably real variability (as wind speed is not the only driver of whitecap fraction, and color of whitecaps depends on things like their size and age). This is discussed a bit in some work the authors cite (e.g. Frouin for whitecap reflectance) and by other authors (e.g. M Anguelova, A. Callaghan). For example the whitecap fraction expression used in 6S (from Koepke 1984, which I think used Monahan?) comes out, I think, quite a lot higher at high wind speeds than more recent measurements. As the present study is looking in large part at radiative transfer simulations this uncertainty doesn't matter so much (as the authors are looking at reflectance changes), but on application in a real retrieval, it's possible that the error from neglecting bubble rafts could be swamped in the error from the whitecap fraction or whitecap reflectance parameterizations, or even from the error on the ancillary wind data ingested. I think this should be acknowledged more directly in the manuscript, but am unsure of the best way to resolve it. One thing to look at would be how large the modelled raft contribution is compared to the modelled whitecap contribution. Addressing this would help to figure out whether retrieval errors can be best reduced by focusing on improving parameterization of whitecap fraction/reflectance, better ancillary wind speed data, or including rafts in models, first.

We thank the reviewer for their comments. These issues have been taken care of through experimentation. We have added further clarifications to the paper. For example, the relative contributions from whitecaps and submerged bubbles are included in the paper and we have added the following discussion:
"Also, this is only a theoretical analysis. While in practical applications, the uncertainties in whitecap estimates (fraction coverage and spectral reflectance, e.g. Frouin et al., 1996; Anguelova et al., 2006) need to be fully considered and incorporated into the analysis. "