[REPLY] We thank the reviewer for her/his excellent comments. When responding to the reviewer’s comment #5, we found that we had made a mistake in the calculations using the full, prognostic WRF-Chem model. During the optical calculations, we had retained the largest two size bins, so the optical properties so calculated were not comparable to the fine mode aerosol, which has a size cut-off of about 2.5 um aerodynamic diameter. This problem was easily corrected by chopping off the two largest size bins in the full WRF-Chem calculations, but only when finding aerosol optical properties. (We did not make this error when evaluating the offline “chemical to optical properties” module.) Some of the numbers associated with the full, prognostic WRF-Chem calculations changed, for example, for the old calculation, the SSA was 0.90 and for the new calculations, the SSA is 0.87. Fortunately, the new calculations strengthen the conclusions expressed in the paper. In particular, the role of BC emissions in causing errors in SSA is now seen more clearly; see comment #5 below. Figure 1 was also changed to show the corrected simulation.

We gently remind the reviewer that the major point of the paper is to evaluate the isolated, offline “chemical to optical properties” module, using measured chemical properties as input to the module. We make a few comments about the full, prognostic WRF-Chem calculations, and for these comments, information about emissions, etc. is important. But it must be emphasized that we are not evaluating all aspects of the full WRF-Chem model. That would be a gigantic job, far beyond the scope of this paper.

Summary

The authors show that discrepancies in aerosol optical properties (B\text{\textsubscript{scat}}, B\text{\textsubscript{abs}}, and SSA) between the WRF-Chem simulation and measurements obtained during the MILAGRO campaign at the T1 site are not attributable to the module in WRF-Chem used to convert aerosol chemical properties to aerosol optical properties. Namely, by imposing speciated mass concentrations using detailed data from the T1 site, the WRFChem module performs within estimated uncertainties. The study is a very nice example of how measurements can be used to support model evaluation, how to consider uncer-
tainties from both the measurements and from the model when analyzing the comparisons, and how it is not enough to condemn models as simply a ‘black box’.

General Comments

The authors show that the WRF-Chem chemistry-optics module is not to blame for the poor agreement shown in Figure 1. They also suggest that emissions input to WRFChem (full version, corresponding to Figure 1) are likely to blame. There are some points related to this that I think the authors should clarify:

1. Please add a reference to the emissions inventories (fossil fuel, biomass burning, biofuel, etc.) you are using as input.

[REPLY] The emissions inventory used for WRF-Chem modeling during MILAGRO has been discussed in great detail in Fast et al. (2009); see point two below.

2. Since you are discussing surface comparisons, can you briefly state how you vertically inject emissions into WRF-Chem and whether this has any bearing on your conclusions?

[REPLY] The vertical injection of emissions was also discussed in Fast et al. (2009). Biomass emissions were distributed uniformly within about 300 meters of the ground. Stack emissions took place at stack height. Because the emissions are so thoroughly discussed in Fast et al., we simply refer the reader to this reference. Here’s the relevant text, placed in the introduction:

Fast et al. (2009) describe the WRF-Chem configuration used for MILAGRO as well as the emission inventories. In regards to the evaluation of the “aerosol chemistry to aerosol optical properties” module, the vertical injection of emissions has no bearing on our conclusions, because the chemical measurements used for this evaluation are not influenced by the emissions assumptions used in the full WRF-Chem model.

3. How do the WRF-Chem diurnally averaged mass concentrations derived from that emissions inventory compare to the observed values presented in Figure 2?

[REPLY] It would be nice to have time series of the various species as calculated by WRF-Chem, but an easier and clearer route to illustrating the effect of these species on optical properties is to show only BC (relevant to absorption at 870 nm) and PM2.5 (relevant to scattering). I have put this information in Table 3, which shows the WRF-Chem derived concentrations, and the measurements, of EC and PM2.5. This table has required additional text in the manuscript, inserted just prior to the “Uncertainties” section.

For the sake of comparison, it is interesting to show the aerosol optical properties for the full, prognostic WRF-Chem run. These are shown in the fourth column of Table 2. The single scattering albedo, for the full 10-day period, is 0.87, as noted above. The prognostic WRF-Chem simulation substantially underpredicts Babs for all time periods, and overpredicts Bscat for the full time period as well as for the showery period. For the clear period, both observed and calculated Bscat values are about the same. We then ask, why is Babs so grossly underpredicted? Table 3 shows PM2.5 and BC concentrations, both measured and as predicted by WRF-Chem. A comparison of the BC concentrations reveals that the amount of BC found in the WRF-Chem simulation is far lower than the measurements; for example, for the full time period, the BC concentration is 0.70 μg/m³ for WRF-Chem, yet the measured value is 1.54 μg/m³. Because BC is a primary emission that is not altered significantly in the atmosphere, we attribute WRF-Chem’s poor simulation of BC to the emissions inventory that does not contain enough BC.

This table also shows that on an overall basis, the prediction of PM2.5 is similar for WRF-Chem (32.7 μg/m³) and the observations (28.9 μg/m³), but major differences occur in the clear and showery period. During the clear period, WRF-Chem significantly underpredicts the PM2.5 mass, and the opposite is true in the showery period. We cannot yet explain this behavior. Because PM2.5 is closely related to the scattering (at 870 nm), when the predicted PM2.5 is too large relative to the observations, the predicted Bscat is similarly too large, and vice versa for the predicted PM2.5, which is too
small. For example, during the showery period, the simulated and observed PM2.5 values are 40.0 µg/m³ and 21.2 µg/m³, respectively, while the simulated and observed Bscat is 56.7 Mm⁻¹ and 28.7 Mm⁻¹, respectively. A doubling of PM2.5 leads to a doubling in the scattering. For the showery period, the discrepancy between predicted and observed PM2.5 significantly influences the scattering, and therefore the value of SSA. If we calculate SSA using the observed PM2.5 (less scattering) in place of the WRF-Chem PM2.5 (more scattering), we find that SSA drops by about 0.09.

To bolster the conjecture that the specified emissions of BC are too low, we start with the chemical concentrations as simulated by WRF-Chem. We make a single change to these concentrations: we replace the simulated BC concentration by the observed concentration of BC. When this new input is fed to the module, the overall SSA value is now 0.78, the same value as the observations. However, during the clear and showery periods, there remain significant differences (0.05 and 0.07) between the observed and calculated SSA values. The various optical properties as simulated by the module, using WRF-Chem predicted chemical concentrations, with the predicted BC replaced by the measured BC, is shown in the fifth column of Table 2.

We also added some text to the conclusions to reflect these findings:

When the observed BC mass concentrations at T1 were substituted into the full simulation (not shown), replacing the simulated values derived from the emissions inventory estimates used as model input, the simulated mean value for SSA decreases to 0.78, which is the same as the observed mean value of 0.78. Moreover, a significant portion of the diurnal variation in SSA was then simulated. This suggests that BC emissions may be the primary cause for the poor SSA simulation by WRF-Chem. However, for the showery case, WRF-Chem overpredicts PM2.5 mass, leading to an overprediction of scattering and therefore SSA. This shows that predicted PM2.5 mass also plays a role in determining aerosol optical properties.

4. If BC emissions are particularly bad, as you say in the Abstract, perhaps we need more information about the original WRF-Chem output than just Figure 1. Table 2, for example, could be expanded by one column to include the mean aerosol optical properties from the full (original) WRF-Chem simulation.

[REPLY] This is a very good idea, and we have added full WRF-Chem derived optical properties to Table 2 (column 4). The implications of mentioning the full, prognostic WRF-Chem run is discussed by the new text listed under point #3 above.

5. On p. 8947, lines 1-5, you say that you ran the full simulation using observed BC mass and that resulting SSA was 0.85 (compared to 0.78 from observation, 0.74 from aerosol module with observed inputs). Can you clarify what this 0.85 implies? Why is it so much higher than 0.74, for example? And why not run the full simulation with all observed mass concentrations to see how this affects the comparisons?

[REPLY] We noted the error in our full, prognostic WRF-Chem calculations above. For the WRF-Chem simulation, the SSA was 0.87 for the corrected calculation (as opposed to 0.90 in the original calculation). When we use the WRF-Chem chemical simulations as input to the module, with the exception of replacing simulated BC with observed BC, the SSA becomes 0.78, the same as the observations. This is discussed in the new text above, under comment #3. The point that BC emissions are primarily responsible for the difference between full, WRF-Chem simulations of SSA and the observations of SSA is now stronger. (Although the fact that both of these SSAs are exactly 0.78 is likely coincidental).

I don’t see exactly how we would run the full simulation with the observed masses. I guess we could use something like 4-D data assimilation to constrain the predicted masses to the observed mass at the points where we have observed masses. But such a run is far outside the scope of this paper.

Specific Comments
1. p. 8232, lines 15-17: Your assumption about the BC refractive index is probably fine,
but to be fair to Bond and Bergstrom (2006), you should add here that they suggest $1.85+0.71i$ as the midpoint of a range of values of refractive index at 550 nm and explicitly say in Section 9.1 that they make ‘no statements about the nature of absorption at ultraviolet and infrared wavelengths’ and in Section 9.2 say ‘we emphasize that it [the refractive index] is likely to be much different’ at infrared wavelengths. You should also add a note to Table 1 saying the same thing about your assumed 870 nm BC refractive index.

[REPLY] We have added these lines to the text:

...specified at 550 nm. We note here, however, that Bond and Bergstrom state that this refractive index may be assumed to be constant across the visible spectral region, extending from 400 nm to 700 nm, but it may be much different at ultraviolet and infrared wavelengths.

We also added a note in Table 1 saying that the BC refractive index is for 550 nm.

2. p. 8934, lines 13-15, Figure 2: How significant are the peaks and valleys in the diurnally averaged chemical concentrations compared to the variability (standard deviation) in the 4-6 day averages?

[REPLY] Hmmm ... Don’t quite understand this comment. I’m not sure what the reviewer means by ‘4-6 day averages’; I assume these averages are over the clear and showery periods, and that the reviewer is asking for standard deviations over each period.

Assuming that I understand the reviewer’s request correctly, here are the standard deviations for selected chemical constituents:

PM2.5: 14.3 (all), 13.7 (clear), 9.1 (showery)
BC: 1.36 (all), 1.39 (clear), 1.24 (showery)

I would say, using a visual assessment, that the peaks and valleys exhibited in the diurnally averaged chemical concentrations (looking specifically at PM2.5) are roughly the same as the standard deviations. But I am not sure what this proves, so I haven’t added any new material to the paper. It has no bearing on the outcome of the paper.

3. p. 8939, line 26: Figure 1 shows SSA, not BC concentration.

[REPLY] taken care of: Figure 1 in the text is now correctly labeled as Figure 2

4. p. 8944, lines 15-26: Nice technique!

[REPLY] thanks!

5. Table 1: Similar to the caveat about the assumed BC refractive index (my comment about p. 8232, lines 15-17), you should also add a note to Table 1 saying that the OC review by Kanakidou et al. 2005 only summarized work for a wavelength range of about 300-800 nm.

[REPLY] added text in the paper that states:

... Kanakidou et al. (2005) report ranges of $n$ extending from about 1.35 to 1.75 (for a wavelength range of about 300 nm to 800 nm), ...

6. Table 2, Figure 1, Figure 6: Add ‘at 870 nm’ somewhere in these.

[REPLY] Done!

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 8927, 2010.