Review of Park et al., Retrieval Algorithm for Aerosol Effective Height from the Geostationary Environment Monitoring Spectrometer (GEMS), submitted to AMT

This paper presents preliminary results of GEMS AEH retrievals and shows the validation results. The performance of the algorithm is highlighted in the form of comparisons to correlative measurements from other sensors such as CALIOP and TROPOMI. The performance metrics are stratified in different ways to demonstrate what input parameters affected the retrievals and so forth. I must admit I had great difficulty reading this paper. The content of the paper is rather simple and easy because the analysis is straightforward. However, the writing style and leaps of faith in the retrievals boggled my mind. The whole Introduction section is so bizarre; the authors could not explain in clear terms why retrieving AEH is important and what work has been done so far by other researchers – basically a literature survey. In its current form, the manuscript cannot be published. Work can be improved and will be useful. So, a rejection with resubmission with better scientific discussion of findings and succinct writing is recommended. Specific issues:

(1) These retrievals are all new and I encourage the authors to show some look-up-table information. Without that information, it is hard to believe the conclusions. For example, AEH retrieval is sensitive to AOD and SSA. How strong is the sensitivity? Without this information, how are we supposed to believe that uncertainty in AEH stems from uncertainties in AOD or SSA or both? I can understand SSA uncertainties could be large but we can retrieve AOD reasonably accurately.

(2) In the algorithm description section, the authors mention that they use surface reflectance climatology but it looks like retrieved surface reflectance is used? Because they attribute part of the AEH uncertainty to surface reflectance uncertainty.

(3) GEMS vs. CALIOP AEH differences are quite remarkable. Sure, the mean bias is small or within a kilometer or so but the scatter plots indicate no correlation whatsoever. By presenting just mean bias metrics and not discussing root mean square error (rmse) and correlation coefficient, the authors are sending the wrong message. All analysis should use rmse or precision (standard deviation of the bias) to fully understand the performance of the algorithm. Standard deviations are shown in some cases but those are too large. Does GEMS have a specification for AEH retrieval? If yes, do these biases, precisions, and uncertainties (rmse) meet those specifications?

(4) Instead of doing gross statistics for land vs. ocean, why not plot bias, precision, uncertainty (rmse) as a function of surface reflectance? Land vs. ocean statistics are fine but show the dependence plot for surface reflectance as well.

(5) What are the reported uncertainties in GEMS AOD and SSA retrievals? Over different surface types and different view conditions (scattering angle or time of the day). How do they relate to AEH retrievals?

(6) Again, not knowing how sensitive AEH is to SSA (show LUTs for us to visually see the dependence), it is hard to tell how the performance of the algorithm would be for dust only, smoke only, smoke mixed with urban pollution etc.

(7) There is a positive bias for dust and negative bias for absorbing aerosol and scattering aerosol. This is not explained well. What constitutes scattering aerosol? Sea salt or sulfates and organics?

(8) How different is the GEMS SSA for dust cases, absorbing aerosol cases, and non-absorbing cases? Can some information be provided?

(9) For some of the data points in the scatter plots between GEMS and CALIOP, it would be nice to show the vertical extinction profiles to see if aerosol layers are continuous or stratified? Maybe, the issue is not with GEMS AEH but with how CALIOP ALH is being derived?

(10) Colors in figures are not consistent. For example, in Figure 9b, red color is used for scattering aerosols (SSA > 0.95) and in Figure 9c, red color is used for dust. Very confusing to interpret the figures when colors are mixed and not consistent

In summary, this paper is not qualified to be published as it is neither written well nor the scientific results good. Nothing wrong with a poor performance of an algorithm/product; but the results are not presented that way. The results are presented as if the retrievals are good, which is not the case. Again, I know this is not an easy product to retrieve. I wish the authors said so at the outset and presented the results as is. The authors also leave out what they plan to do to improve the product performance. There is a lot of work yet to be done for this paper to be published.