Review of "Aerosol Scattering Effects on Water Vapor Retrievals over the Los Angeles Basin" by Zeng et al.

This paper uses retrievals of H_2O SCDs from CLARS observations to demonstrate that standard deviations of retrieved H_2O SCDs amongst 15 H_2O bands are due to the effects of scattering aerosol. This fact is established by the calculations.

The results in Figure 3 are very reasonable. The panel 3a (SVO observing path from Mt. Wilson to the sun) H_2O SCDs have little wavelength dependence since little aerosol impacts these observations. The panel 3b (LABS LA basin observing path) H_2O SCDs display much wavelength dependence, since aerosol is not included in the retrieval, and aerosol is an important contributor to the total optical depths. The retrieval of the LABS data therefore needs to add additional H_2O in lieu of aerosol that is present in the LA basin. Panel 3d indicates that the AERONET-Caltech aerosol increases during the day, and therefore the H_2O SCD standard deviations in panel 3c also increase during the day.

This paper, however, is very problematic since the methodology does not go far enough in its analyses. Yes, the data in Figure 3 is consistent with the fact that aerosol impacts the total optical depths in the H_2O bands. This is expected, and by itself is not a sufficient reason for publication. The current paper does not demonstrate that accurate optical properties (e.g. AODs) can be readily retrieved from the observations.

The paper should be published after additional calculations are carried out by the authors.

Major comments

The following suggested calculations would bring the paper to a level of completion that fully warrants publication:

(a) The retrieval program retrieves H_2O SCDs and aerosol properties (e.g. vertical AOD) simultaneously as a function of wavelength. Representative AERONET-MERRA-GOCART SSA and g could be specified (fixed) in a daily basis in the forward model of the retrieval.

(b) Compare the retrieved aerosol properties (e.g. AOD) to those derived from a combined analysis of AERONET, MERRA, and GOCART data,

(c) The retrieval program retrieves H_2O SCDs with specified AERONET-MERRA-GOCART wavelength dependent AOD, SSA, and g

 (d) Demonstrate that RT forward model calculations including the retrieved H_2O and aerosol properties (from (a)) reduce the scatter in a Figure 3b type graph, and

 (e) Demonstrate that RT forward model calculations including the retrieved H_2O (from (a)) and AERONET-MERRA-GOCART aerosol properties reduce the scatter in a Figure 3b type graph.

The sentence in the Abstract "The understanding of aerosol scattering effects on H_2O retrievals provides a sensitive way to quantify the effect of aerosol scattering on greenhouse gas retrievals …" indicates that this fact is established by the work in this paper. The phrase "provides a

sensitive way" is not demonstrated by the work in this paper. This sentence needs to be removed from the abstract. The language on page 2, lines 19-20 is appropriate and can be retained. The language on Page 4, line 29 "shows the potential" is appropriate. The final sentence (page 10, lines 23-26) with the phrases "evidence justify our approach" and "providing a sensitive way" is not demonstrated by the paper's calculations.

Page 3, line 12. Is wavelength dependent surface reflection included in the CLARS-FTS GFIT retrieval algorithm?

Page 3, line 12. Why is wavelength dependent aerosol AOD (and possibly mean SSA and g) not included in the retrieval?

Page 3, line 21. How is Figure 2 constructed? The Figure caption refers to "normalized radiance", yet it is not stated in the paper how the normalized radiance is calculated. Please do so. Is aerosol included in the RT model calculations that are presented in Figure 2?

Page 3, lines 24-25. A full information analysis would follow the Rodgers methodology and include (a) calculations using a retrieval state vector that includes both H_2O and aerosol (with little influence by the *a priori* aerosol) and (b) calculations using a state vector that includes H_2O and aerosol heavily constrained by the AERONET-MERRA-GOCART data.

Page 6 and Figure 6. It is requested that panel (a) also be presented with the means included. The results in panels (a) and (c) are opposite to what is expected. If one adds the AERONET data to the forward model RT calculation, I would expect that the scaling factors would be closer to unity (and closer to 0 when the means are subtracted from the scaling factors) than for the case when no aerosol (panel (c)) is included in the forward model RT calculations. Yet the opposite is apparent. Please clarify.

Page 6 and Figure 6. If the standard deviations are not normalized, what does panel 6d look like? Again, the Aerosol-Free and Clear day curves seem to indicate that less *a priori* information (and/or a less complete inclusion of all contributors to the forward model) produces a better result. Please clarify.

<u>Minor comments</u>

Page 3, line 27. It is important to mention that aerosol is also included in the forward model RT.

Page 4, line 1 Briefy mention how IC is calculated.

Page 4, line 3. State which variables are retrieved.

Page 4, lines 17-18. Indicate (in %) the representative "small differences" and "larger variation" values.

Page 5, line 11. State the wavelength range "wavelengths (i.e. from 1288 nm to 2190 mn))"

Page 5, line 12. The sentence is not clear. Is the "AOD data" a vertical optical depth, and the AOP is simply this value scaled by a SZA and CLARS viewing angle geometric airmass factor?

Page 5, line 21. Why should the PBLH be similar throughout the year? Is the PBLH information available from nearby airport radiosonde temperature-pressure profiles or from any camera images taken at CLARS? On page 6 the PBLH data is from spring 2010, yet the observations are in winter-spring and summer-autumn 2013. Explain why (and if) the PBLH values are so uniform. On page 7 (line 1) it is stated that the PBLH is an important parameter. Should it not be included in the retrieval state vector?

Page 5, line 23. The "other factors" are not discussed in the text. What are the "other factors"?

Page 6, line 13. Why is HITRAN 2008 (instead of HITRAN 2012) used in the RT calculations?

Page 6, line 14. Is the surface albedo of 0.23 really wavelength independent from 1288 nm to 2190 nm?

Page 8, line 27. Is annual average MERRA data or month-specific MERRA data used for each individual day of the 68 day set?