Water vapor retrieval from OMI visible spectra

H. Wang\textsuperscript{1}, X. Liu\textsuperscript{1}, K. Chance\textsuperscript{1}, G. González Abad\textsuperscript{1}, and C. Chan Miller\textsuperscript{2}

\textsuperscript{1}Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
\textsuperscript{2}Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA

Correspondence to: H. Wang (hwang@cfa.harvard.edu)

Received: 2 December 2013 – Published in Atmos. Meas. Tech. Discuss.: 22 January 2014
Revised: 21 April 2014 – Accepted: 23 May 2014 – Published: 30 June 2014

Abstract. There are distinct spectral features of water vapor in the wavelength range covered by the Ozone Monitoring Instrument (OMI) visible channel. Although these features are much weaker than those at longer wavelengths, they can be exploited to retrieve useful information about water vapor. They have an advantage in that their small optical depth leads to fairly simple interpretation as measurements of the total water vapor column density. We have used the Smithsonian Astrophysical Observatory (SAO) OMI operational retrieval algorithm to derive the slant column density (SCD) of water vapor using the 430–480 nm spectral region after extensive optimization. We convert from SCD to vertical column density (VCD) using the air mass factor (AMF), which is calculated using look-up tables of scattering weights and assimilated water vapor profiles. Our Level 2 product includes not only water vapor VCD but also the associated scattering weights and AMF. In the tropics, our standard water vapor product has a median SCD of $1.3 \times 10^{23}$ molecules cm$^{-2}$ and a median relative uncertainty of about 11\%, about a factor of 2 better than that from a similar OMI algorithm that uses a narrower retrieval window. The corresponding median VCD is about $1.2 \times 10^{23}$ molecules cm$^{-2}$. We have examined the sensitivities of SCD and AMF to various parameters and compared our results with those from the GlobVapour product, the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Aerosol Robotic NETwork (AERONET).

1 Introduction

Water vapor is one of the key factors for weather. It is also the most abundant greenhouse gas in the atmosphere. It can provide strong feedback directly through amplification of global warming associated with other greenhouse gases and indirectly through formation of clouds. Water vapor participates in many photochemical reactions, such as the reaction with O($^1$D) to produce OH radicals, which control the oxidation capacity of the atmosphere. It is therefore also important for atmospheric chemistry. Unlike other long-lived greenhouse gases, the short-lived water vapor exhibits large spatial and temporal variability. Monitoring the distribution, variability and long-term changes in water vapor is critical for understanding the hydrological cycle, the earth radiative budget and climate change.

Water vapor has long been observed using a variety of platforms and methods, including measurements made from surface stations, balloons, aircrafts and satellites using in situ or remote sensing techniques. Satellite observations of water vapor have used microwave and radio wave (e.g., Advanced Microwave Sounding Unit (AMSU), Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E), Special Sensor Microwave/Imager (SSM/I), Special Sensor Microwave Imager/Sounder (SSMI/S) (Ferraro et al., 2005; Boukabara et al., 2010), Global Navigation Satellite System (GNSS) (Lee et al., 2013)), thermal infrared (e.g., Moderate Resolution Imaging Spectrometer (MODIS) (King et al., 2003), Tropospheric Emission Spectrometer (TES) (Worden et al., 2012), Atmospheric Infrared Sounder (AIRS) (Aumann et al., 2003), Infrared Atmospheric Sounding Interferometer (IASI) (Schneider and Hase, 2011)), and near-infrared and visible sensors (e.g., MODIS (King et al., 2003), Global Ozone Monitoring Experiment (GOME) (Noël et al., 2002; Wagner et al., 2003), GOME-2 (Grossi et al., 2014), Scanning Imaging Absorption spectrometer for Atmospheric CHartographY (SCIAMACHY) (Noël et al., 2004), Medium Resolution Imaging Spectrometer (MERIS) (Lindstrot et al., 2012)).
Figure 1. Reference spectra used in the standard operational water vapor retrieval. The spectra have been scaled for presentation purposes. The black lines are those listed in Table 1. The red lines are the black lines convolved with the OMI slit function.

Wagner et al. (2013) demonstrated the feasibility of water vapor retrieval in the blue spectral range using GOME-2 and Ozone Monitoring Instrument (OMI) measurements. They pointed out that the advantages of this spectral range include more consistent retrievals across the globe due to more uniform surface albedo, especially between land and ocean, increased sensitivity to near-surface layer due to higher surface albedo over the oceans than for longer wavelengths, less saturation of signal due to weaker water vapor absorption, applicability to sensors that do not cover longer wavelengths, and daily global coverage over a long period of time. Wagner et al. (2013) derived OMI water vapor slant column densities (SCDs). We have independently derived OMI water vapor SCDs and converted them to vertical column densities (VCDs) using the Smithsonian Astrophysical Observatory (SAO) operational retrieval algorithm. In this paper, we will present our SCD retrievals, VCD calculations, sensitivity studies and initial validation results.

2 Data processing

2.1 OMI instrument and OMI data

OMI is a joint Dutch–Finnish instrument onboard the NASA EOS-Aura satellite, which was launched on 15 July 2004 into a sun-synchronous orbit with an ascending node Equator crossing time of around 13:45 LT and an orbital period of about 100 min (Schoeberl et al., 2006). It is a nadir-viewing push-broom ultraviolet/visible (UV/VIS) imaging spectrometer with three channels – the UV1 (264–311 nm), UV2 (307–383 nm) and VIS (349–504 nm) – at 0.42–0.63 nm spectral resolution (Levelt et al., 2006). For the visible channel, the 2600 km OMI cross-track swath usually provides a nominal spatial resolution between 13 km × 24 km at nadir and 26 km × 135 km at the edge. The entire globe is covered by 14–15 orbits each day. Solar irradiance measurements are performed daily.

We use Version 3 Level 1B OMI visible spectra to derive water vapor SCD, and Version 3 Level 2 OMI cloud pressure and cloud fraction product (OMCLD02) downloaded from disc.sci.gsfc.gov/Aura/data-holdings/OMI/ for air mass factor (AMF) calculation.

2.2 Slant column retrieval

2.2.1 Standard retrieval

We determine the slant column of water vapor by directly fitting the OMI spectra following the method described in Chance (1998). The method is also presented in detail in González Abad et al. (2014). In this paper, we only provide a brief description.

Wavelength calibration is performed using cross-correlation through spectral shift (Caspar and Chance, 1997) with a high-resolution solar reference spectrum (Chance and Kurucz, 2010). To reduce noise, we use the leading principle component derived from OMI solar spectra as the measured solar spectrum. The slant column abundance that minimizes the difference between the measured and calculated radiance is retrieved using a non-linear least squares inversion method by Lindström and Wedin (1988). We use a spectral window from 430 nm to 480 nm for our standard water vapor retrieval as it leads to the smallest retrieval uncertainty (Sect. 2.2.2). A third-order polynomial is fitted for both the baseline and the scaling factor to account for broadband spectral features. Common mode and under-sampling spectra (Chance et al., 2005) are derived on-line and applied to the fitting. The retrieval takes into consideration water vapor, ozone, nitrogen dioxide, oxygen collision complex, liquid water, glyoxal, the ring effect (Chance and Spurr, 1997) and the liquid water ring effect. The molecular reference spectra used in our standard retrieval are listed in Table 1 and plotted in Fig. 1. There are distinct spectral signatures of these molecules in our retrieval window. The reference spectra
are interpolated onto common calibrated radiance grid and convolved with pre-determined instrument slit function (Dirksen et al., 2006) during the fitting (Fig. 1). Ozone, nitrogen dioxide, water vapor and glyoxal are corrected for the solar \( I_0 \) effect (Aliwell et al., 2002).

Figure 2 shows our standard retrieval result for the SCD and the associated absolute and relative uncertainties for 14 July 2005 (orbits 5297–5311). As expected, the global pattern shows more water vapor in the Intertropical Convergence Zone (ITCZ) and mid-latitude weather systems. There are some stripes along the swaths (Veihelmann and Kleipool, 2006) as we have not applied our post-processing routine to remove them for this plot. The stripes are mainly caused by OMI systematic measurement errors and are common to most OMI Level 2 products. In the tropics (30° S–30° N), the median of SCD is 1.32 \( \times 10^{23} \) molecules cm\(^{-2}\), the median of fitting uncertainty is 1.2 \( \times 10^{22} \) molecules cm\(^{-2}\), the median of relative uncertainty (fitting uncertainty/SCD) is 11 % and the median of the fitting root mean square (rms) ratio to the radiance is 9.2 \( \times 10^{-4} \). Areas with larger SCD generally have smaller uncertainties.

The 25 % and 75 % percentiles of our fitting uncertainties in the tropics are 1.0 \( \times 10^{22} \) and 1.7 \( \times 10^{22} \) molecules cm\(^{-2}\), respectively. In comparison, using a shorter retrieval window of 430–450 nm, Wagner et al. (2013) obtained typical SCD uncertainties of 3–5 \( \times 10^{22} \) molecules cm\(^{-2}\) for OMI and 1–2.5 \( \times 10^{22} \) molecules cm\(^{-2}\) for GOME-2. The uncertainty of our standard water vapor SCD is therefore smaller than the Wagner et al. (2013) OMI result and similar to the Wagner et al. (2013) GOME-2 result.

Figure 3 shows examples of our spectral fitting for two pixels from orbit 5306 in July 2005. The left column is for a pixel at 1.75° S and 34.6° W in the Atlantic Ocean, and the right column is for a pixel at 47.75° N and 53.4° W at the Atlantic coast of North America. The retrieved water vapor SCDs are (1.23 \( \times 0.12 \)) \( \times 10^{23} \) and (1.75 \( \times 0.07 \)) \( \times 10^{23} \) molecules cm\(^{-2}\), respectively. The corresponding rms values are 1.1 \( \times 10^{-3} \) and 4.0 \( \times 10^{-4} \), respectively. The panels in the top row show that the fitted spectra (red) closely track the measured spectra (black). The panels in the second row show that the fitting residuals appear random except for two minor noise spikes in the right-hand spectrum. The next four rows show the reference spectra of important molecules (water vapor, liquid water, nitrogen dioxide and ozone) scaled by their corresponding fitted SCDs (black) and added to the fitting residuals in the second row (red). In both cases, the water vapor spectral signature within the fitting window is stronger than the fitting residuals. Consistent with the expectation that there is less liquid water, more NO\(_2\) and more O\(_3\) in the mid-latitude coastal area than in the tropical open ocean, the right panels show that the liquid water signal is weaker and the NO\(_2\) and O\(_3\) signals are stronger than the residual, while the left panels show the opposite.

**Figure 2.** OMI water vapor (top) SCD, (middle) SCD uncertainty and (bottom) SCD relative uncertainty for 14 July 2005 from the standard retrieval.

### 2.2.2 Sensitivity studies

We have investigated the sensitivity of OMI water vapor SCD with respect to the retrieval window. Selected examples are listed in Table 2. We vary the retrieval window while keeping everything else the same. All the retrieval windows in Table 2 include the water vapor feature at about 442 nm. The 438–450 nm and the standard 430–480 nm window also include the weaker water vapor feature at about 470 nm, and the longest window (430–495 nm) includes an additional weaker feature at about 485 nm. The median SCDs and uncertainties within 30° N–30° S for 14 July 2005 are listed in Table 2. The standard window leads to the smallest uncertainty (1.2 \( \times 10^{22} \) molecules cm\(^{-2}\)). The uncertainties for the 20 nm and 30 nm windows are about 100 % and 67 % larger, respectively. The uncertainty for the 65 nm window is about
Table 1. Reference spectrum used in standard retrieval.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>T (K)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water vapor (H$_2$O)</td>
<td>280</td>
<td>Rothman et al. (2009)</td>
</tr>
<tr>
<td>Ozone (O$_3$)</td>
<td>228</td>
<td>Brion et al. (1993)</td>
</tr>
<tr>
<td>Nitrogen dioxide (NO$_2$)</td>
<td>220</td>
<td>Vandeaele et al. (1998)</td>
</tr>
<tr>
<td>Oxygen collision complex (O$_2$-O$_2$)</td>
<td>294</td>
<td><a href="http://spectrolab.aeronomie.be/o2.htm">http://spectrolab.aeronomie.be/o2.htm</a></td>
</tr>
<tr>
<td>Pure liquid water (H$_2$O)</td>
<td>–</td>
<td>Pope and Fry (1997)</td>
</tr>
<tr>
<td>Glyoxal (C$_2$H$_2$O$_2$)</td>
<td>296</td>
<td>Volkamer et al. (2005)</td>
</tr>
<tr>
<td>Ring and water ring</td>
<td>–</td>
<td>Chance and Spurr (1997)</td>
</tr>
</tbody>
</table>

Figure 3. Spectral fitting results for (left) a pixel in the Atlantic Ocean and (right) a pixel near the Atlantic coast of North America. The first row shows the fitted (red) and measured (black) spectra. The second row shows the fitting residuals. The third to sixth rows show the reference spectra of H$_2$O, liquid water, NO$_2$ and O$_3$ scaled by the fitted slant columns (black) and added to the fitting residuals (red).

25\% larger. The median SCD decreases from $1.47 \times 10^{23}$ to $1.23 \times 10^{23}$ molecules cm$^{-2}$ as the retrieval window length increases.

We have performed additional sensitivity studies, shown in Table 3, by excluding the interfering molecules, changing the reference spectra and changing the order of closure polynomials. In these experiments, everything else is kept the same as in the standard retrieval. In Table 3, we list the median statistics and the number of negative retrievals for water vapor between 30° S and 30° N for 14 July 2005.

Exclusion of O$_3$, O$_2$–O$_2$, NO$_2$ or liquid water leads to significant (10–30\%) reduction of the retrieved water vapor SCDs and large increase of the number of negative retrievals, though the fitting uncertainties and rms remain at the same level. The most severe change is associated with liquid water, followed by NO$_2$, O$_2$–O$_2$ and O$_3$. Exclusion of C$_2$H$_2$O$_2$ leads to only about 1\% increase of water vapor SCD. Without liquid water, the medium water vapor SCD decreases by about 32\% from $1.32 \times 10^{23}$ to $0.90 \times 10^{23}$ molecules cm$^{-2}$, and the number of negative retrievals increases from 1935 to 50 216. It should be noted that such a strong sensitivity to liquid water is for the standard long retrieval window of 430–480 nm. For the shorter window of 432–462 nm, the difference in the median SCD with and without liquid water is only about 4\%, which is substantially smaller than the median relative uncertainty.

As a by-product of our standard water vapor retrieval, the top panel of Fig. 4 shows the retrieved liquid water on 14 July 2005. Although the retrieval is not optimized for liquid water, areas in the oceans, seas, gulfs and so on are highlighted. Not all liquid water bodies are highlighted to the same extent. Comparison between the top and middle panels of Fig. 4 shows that some liquid water surfaces are shielded by clouds. The bottom panel of Fig. 4 shows the retrieved water vapor SCD without considering liquid water. Compared to the standard retrieval shown in the top panel of Fig. 2, the SCDs here are apparently smaller, especially over the areas with liquid water where many negative values (plotted as blanks) are retrieved.

To further investigate the influence of liquid water, we examine the common mode in Fig. 5 for Orbit 5311, which is dominated by ocean and Orbit 5304, which is dominated by...
land. The common mode for each orbit is derived from the average of the fitting residuals. The fitting program then uses the derived common mode as a reference spectrum for the final retrieval. Since common mode is fitted, the change in the median fitting rms in Table 3 is small. Further exclusion of the common mode from the retrieval without liquid water will lead to an increase of the median rms for the day from 9.5 × 10⁻⁴ to 1.13 × 10⁻³. The common modes for the retrieval without liquid water are shown in the second row. There are apparent spectral structures in the common mode for the ocean-dominated Orbit 5311. In comparison, the common appears more random for the land-dominated Orbit 5304. The median rms in this case is 9.7 × 10⁻⁴ for Orbit 5311 and 8.5 × 10⁻⁴ for Orbit 5304. When liquid water is included in the retrieval, the bottom row shows that the spectral structures of the common mode for Orbit 5311 are reduced while those for Orbit 5304 are little affected. The median rms for the standard retrieval is 9.4 × 10⁻⁴ for Orbit 5311 and 8.4 × 10⁻⁴ for Orbit 5304. The remaining structures in the common mode of the standard retrieval over the ocean probably suggest errors in the liquid water reference spectrum. This will be investigated further in the future.

After obtaining new reference spectra for water vapor and oxygen collision complex, we have tested the sensitivity of our standard retrieval with respect to them. Switching from HITRAN 2008 (Rothman et al., 2009) to HITRAN 2012 (Rothman et al., 2013) water vapor reference makes the median SCD about 6 % lower than that of the standard retrieval. In comparison, the median relative uncertainty of the standard retrieval is about 11 %. Switching to the Thalman and Volkamer (2013) O₂–O₂ reference spectrum gives almost the same result as the standard retrieval, so does switching from a third-order to a fifth-order closure polynomial for the baseline and scaling factor. The median of SCD retrieved using water vapor reference spectrum at 0.7 atm and 265 K is about 2 % lower than the standard result, and that using water vapor reference at 1.0 atm and 288 K is about 2 % higher. Due to the small changes in the SCD statistics, we have not updated the standard retrieval with the new reference spectra.

### 3 Vertical column calculation

#### 3.1 AMF calculation

To derive the VCD, we divide SCD by AMF, i.e., \( \text{VCD} = \frac{\text{SCD}}{\text{AMF}} \). The AMF can be expressed for optically thin absorption as the product of the vertical integral of the shape factor and the scattering weight. The shape factor accounts for the normalized vertical profile of water vapor (Palmer et al., 2001). Averaging kernels can be derived as

<table>
<thead>
<tr>
<th>Window length (nm)</th>
<th>Retrieval window (nm)</th>
<th>Median SCD (molecule cm⁻²)</th>
<th>Median uncertainty (molecule cm⁻²)</th>
<th>Median relative uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>[435, 455]</td>
<td>1.47 × 10²³</td>
<td>2.4 × 10²²</td>
<td>0.19</td>
</tr>
<tr>
<td>30</td>
<td>[432, 462]</td>
<td>1.43 × 10²³</td>
<td>2.0 × 10²²</td>
<td>0.17</td>
</tr>
<tr>
<td>40</td>
<td>[438, 478]</td>
<td>1.35 × 10²³</td>
<td>1.6 × 10²²</td>
<td>0.15</td>
</tr>
<tr>
<td>50 (standard)</td>
<td>[430, 480]</td>
<td>1.32 × 10²³</td>
<td>1.2 × 10²²</td>
<td>0.11</td>
</tr>
<tr>
<td>65</td>
<td>[430, 495]</td>
<td>1.23 × 10²³</td>
<td>1.5 × 10²²</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 3. Miscellaneous sensitivity studies.

<table>
<thead>
<tr>
<th>Description</th>
<th>Median SCD (molecule cm⁻²)</th>
<th>Median uncertainty (molecule cm⁻²)</th>
<th>Median rms</th>
<th>Number of negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1.32 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>1935</td>
</tr>
<tr>
<td>Without O₃</td>
<td>1.19 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.3 × 10⁻⁴</td>
<td>7234</td>
</tr>
<tr>
<td>Without O₂–O₂</td>
<td>1.18 × 10²³</td>
<td>1.3 × 10²²</td>
<td>9.9 × 10⁻⁴</td>
<td>5076</td>
</tr>
<tr>
<td>Without NO₂</td>
<td>1.05 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.3 × 10⁻⁴</td>
<td>15666</td>
</tr>
<tr>
<td>Without liquid water</td>
<td>0.90 × 10²³</td>
<td>1.1 × 10²²</td>
<td>9.5 × 10⁻⁴</td>
<td>50216</td>
</tr>
<tr>
<td>Without C₂H₂O₂</td>
<td>1.34 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>1780</td>
</tr>
<tr>
<td>Switch to fifth-order polynomial</td>
<td>1.32 × 10²³</td>
<td>1.3 × 10²²</td>
<td>9.0 × 10⁻⁴</td>
<td>2262</td>
</tr>
<tr>
<td>Switch reference H₂O to 0.7 atm and 265 K</td>
<td>1.29 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>1992</td>
</tr>
<tr>
<td>Switch reference H₂O to 1.0 atm and 288 K</td>
<td>1.34 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>1918</td>
</tr>
<tr>
<td>Switch to Thalman and Volkamer (2013) O₂–O₂</td>
<td>1.24 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>1816</td>
</tr>
<tr>
<td>Switch to Thalman and Volkamer (2013) O₂–O₂</td>
<td>1.31 × 10²³</td>
<td>1.2 × 10²²</td>
<td>9.2 × 10⁻⁴</td>
<td>2185</td>
</tr>
</tbody>
</table>
Figure 4. The top panel shows the liquid water index from a by-product of our standard water vapor retrieval. The middle panel shows the cloud fraction from the OMCLDO2 product. The bottom panel shows the water vapor SCD from a sensitivity study where liquid water is excluded from the water vapor retrieval. All results are for 14 July 2005.

Figure 5. The top row shows the OMI VCDs for Orbit 5311, which is mostly over the ocean, and for Orbit 5304, which is mostly over the land. The middle row shows the common modes for the two orbits when liquid water is excluded from the standard retrieval. The bottom row shows the common modes for the two orbits from the standard retrieval, which includes liquid water.

the ratio of the scattering weights to the AMFs (Eskes and Boersma, 2003). Our Level 2 product provides VCDs with the scattering weights and AMFs for comparison with or assimilation into models. More details about the AMF calculation for our operational retrieval algorithm can be found in González Abad et al. (2014).

The a priori vertical profiles of water vapor for the shape factor are from the monthly mean early afternoon GEOS-5 data assimilation product. They are generated at the Global Modeling and Assimilation Office (GMAO) and re-gridded to 2° latitude × 2.5° longitude × 47 layer resolution (for GEOS-Chem simulations). We use the monthly mean profiles in our operational retrieval to avoid the need of obtaining near-real-time water vapor assimilation product. The retrieved water vapor VCDs can be easily adjusted using the provided scattering weights when water vapor profiles of higher spatial and temporal resolution are used. As a representative example, the left panel of Fig. 6 shows the monthly and zonal mean water vapor profile at 10° N in July 2007. It can be seen that water vapor is highly concentrated near the surface where the e-folding scale height is approximately 4 km.

The scattering weights are calculated using the VLIDORT radiative transfer model (Spurr, 2006). To speed up the computation, we prepare a scattering weight look-up table based on the surface albedo, observational geometry, surface height and cloud height. For partly cloudy scenes, the scattering weight is approximated as the radiative cloud fraction (φ).
The left panel of Fig. 6 shows a representative water vapor vertical profile in the tropics. The dotted line indicates the height of the 800 mb level. The right panel shows the scattering weight for a (solid line) clear and a (dashed) cloudy atmosphere where the modeled Lambertian cloud surface is at the 800 mb level.

The effective cloud fraction and cloud pressure from the Version 3 Level 2 OMCLDO2 product, which is derived using the O2–O2 absorption band at about 477 nm (Stammes and Noordhoek, 2002; Acaretta et al., 2004; Stammes et al., 2008). In the cloud algorithm, a cloud is represented by a Lambertian reflector with an albedo of 0.8. Consequently, a thin cloud that fully covers an OMI pixel is represented by a small effective cloud fraction. In addition, the retrieved cloud height is mostly inside the cloud. To keep consistency with the OMCLDO2 product, we also model a cloud as a Lambertian surface with an albedo of 0.8.

As an example, the right panel of Fig. 6 shows the scattering weight for a clear (cloud fraction = 0, solid line) and a cloudy (cloud fraction = 1 at 800 mb, which is at a height of about 1.5 km, dashed line) scene under typical conditions. For a clear atmosphere, the scattering weight decreases toward the surface where most of the water vapor resides. For a cloudy atmosphere, the scattering weight shows a jump at the cloud level where the sensitivity immediately above increases due to enhanced multiple scattering and that below drops to zero due to cloud shielding.

### 3.2 AMF Sensitivity

Since errors in AMF affect the quality of VCD, we investigate the sensitivity of the AMF with respect to wavelength, solar zenith angle, surface albedo and cloud pressure in Fig. 7. As a reference, we use a wavelength of 442 nm, surface albedo of 0.05, viewing zenith angle of 0°, solar zenith angle of 30° and surface height of 0 km. The top three panels of Fig. 7 correspond to a clear atmosphere, and the bottom panel corresponds to a cloudy atmosphere with cloud fraction of 1. We vary the parameters of interest one at a time to examine the AMF sensitivity.

The top panel of Fig. 7 shows that AMF is almost insensitive to wavelength. There is only about 1% change over the 430–480 nm range. To speed up computation in our operational retrieval, we use the AMF at 442 nm, which is within the strongest water vapor band in the 430–480 nm retrieval window. The second panel of Fig. 7 shows that the AMF increases from 1.25 to 1.85 as the solar zenith angle increases from 0° to 80°. Since the viewing geometries of satellite observations are precisely determined, errors due to this source can be neglected.

The third panel of Fig. 7 shows that the AMF is quite sensitive to surface albedo. In the blue wavelength region, surface albedo is usually in the range of 0.05–0.15 (Koelmeijer et al., 2003), which corresponds to the steepest slope of the AMF versus albedo curve. For the standard retrieval, we use an updated version of the OMI Lambert equivalent surface reflectance climatology – OMLER (0.5° × 0.5°, http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omler_v003.shtml) (Kleipool et al., 2008). Due to the spectral dependence, spatial heterogeneity and temporal variability, surface albedo estimates can have significant error. Kleipool et al. (2008) showed that the estimated absolute uncertainties were 0.01–0.02 in the blue spectral range. Figure 7 suggests that an increase of surface albedo by 0.02 can result in an increase of AMF by about
Aerosols influence atmospheric scattering and therefore AMF. There are different types of aerosols, and their distributions are highly variable. This can potentially introduce significant error in AMF estimation. However, since the cloud product that we use does not consider aerosols, any effect associated with aerosols is aliased into the cloud information. To be consistent, we do not consider aerosols in our radiative transfer calculation in this paper. In the future, we will perform additional studies to better understand the influence of aerosols on our retrieval.

4 Validation

In this section, we present our initial data validation results. A comprehensive data validation will be performed later. In this paper, we compare our VCDs with the MODIS near-IR data, the GlobVapour MERIS+SSM/I combined data and the AERONET ground-based measurements.

The MODIS near-IR total precipitable water product (Gao and Kaufman, 2003) is derived using the ratios of water vapor absorbing channels (0.905, 0.936 and 0.94 µm) and atmospheric window channels (0.865 and 1.24 µm) in the near-IR. The retrieval algorithm relies on observations of water vapor attenuation of reflected sunlight. Therefore, results only exist for reflective surfaces in the near-IR. The errors are typically about 5–10%, with greater errors over dark surfaces and under hazy conditions. Consequently, the data quality is generally better over the land than over the ocean. In this paper, we use the Level 3 monthly 1° × 1° data from the Aqua platform (MYD08_M3) (ladsweb.nascom.nasa.gov/data/). Aqua is about 15 min ahead of OMI’s host satellite Aura in the “A-train” constellation. Wang et al. (2007) found significant diurnal cycles of precipitable water that vary with region and season. The closeness in local time of observation between OMI and MODIS is nice for comparison.

The top panels of Fig. 8 show our results of the monthly mean 1° × 1° water vapor VCDs derived from the standard retrieval for January and July 2006. For easy comparison with MODIS, we have converted the VCDs from molecule cm$^{-2}$ to (precipitable) cm using a multiplication factor of 2.98904 × 10$^{-23}$ and indicated both units on the color bars of the figure. The gridded OMI data are calculated using the weighted average of the retrieved VCDs whose cloud fractions are less than 0.25. The weight is assigned according to both the area of the ground pixel within the grid box and the fitting uncertainty. The number of data points for a grid box ranges from 0 to about 900 with a median of about 100. The stripes in daily maps (Fig. 2) are averaged out here. The corresponding MODIS results are shown in the second row of Fig. 8. There are lots of missing data (blank areas) in the MODIS maps mainly due to the low near-IR surface reflectivity over the ocean (except under sun glint condition). Both OMI and MODIS show the seasonal shift of the Intertropical Convergence Zone (ITCZ). The OMI–MODIS
The first row shows the monthly mean $1^\circ \times 1^\circ$ OMI water vapor VCDs derived from our standard retrieval for January and July 2006. For easy comparison with the MODIS near-IR total water vapor column in the second row, we have converted the OMI VCDs into precipitable water (cm) and indicated both units on the color bars. The third row shows the OMI–MODIS difference maps. The difference panels are shown in the bottom row of Fig. 8. The difference over the ocean is larger than that over the land. Since MODIS data are most useful over the land, we will focus on the land for subsequent comparison.

The joint probability density distributions of MODIS versus OMI data over land for January and July 2006 are shown in the top row of Fig. 9. We have also indicated the regression lines (solid) and 1:1 lines (dashed) in the plot. The linear correlation coefficients are 0.97 and 0.93 for January and July, respectively. For January 2006, the mean of OMI–MODIS is $-0.06$ cm and the standard deviation is 0.36 cm. For July 2006, the mean of OMI–MODIS is $-0.18$ cm and the standard deviation is 0.30 cm. Figure 9 shows that the range of the data expands, and the mean shifts to higher values from January to July. More than 80% of the data over land have water vapor less than 3 cm. For this subset, the average of OMI is lower than that of MODIS by 0.16 cm in January and by 0.08 cm in July.

The GlobVapour project supported by the European Space Agency (ESA) Data User Element (DUE) program has generated a combined data product from MERIS and SSM/I for 2003–2008 (www.globvapour.info). This product is for observations made over land with MERIS at about 10 a.m. and over ocean with SSM/I at about 6 to 7 a.m. The combination of the two makes an easy-to-use validated global data set of more uniform quality. In Fig. 10, we compare our OMI results with the monthly mean $0.5^\circ \times 0.5^\circ$ GlobVapour combined MERIS+SSM/I data for January and July of 2006. The equation corresponding to the regression line is indicated in each panel.
The first row shows the monthly mean 0.5° × 0.5° OMI water vapor VCDs derived from our standard retrieval for January and July 2006. The second row shows the corresponding maps for the GlobVapour combined MERIS+SSM/I product. The third row shows the OMI–GlobVapour difference maps.

apparently agree better with GlobVapour than with MODIS over the ocean (bottom panel of Fig. 10). The absolute difference between OMI and GlobVapour is also smaller over most land areas, although it is larger in certain cases, such as eastern China and India in July and northern South America and southern Africa in January.

The joint probability density distributions of GlobVapour versus OMI data for the overall, land and ocean area are shown in the bottom three rows of Fig. 9. On global scale, the linear correlation coefficients between OMI and GlobVapour are 0.94 for both January and July of 2006. The mean of the OMI–GlobVapour difference is −0.40 cm in January and −0.30 cm in July, with a standard deviation of 0.53 cm and 0.50 cm, respectively. Over the land, the linear correlation coefficients are 0.97 for January and 0.93 for July. The mean of the OMI–MERIS difference is 0.02 cm in January and −0.05 cm in July, with a standard deviation of 0.39 cm and 0.50 cm, respectively. The linear regression line is quite close to the 1 : 1 line for the land. Over the ocean, the linear correlation coefficients are 0.95 for January and 0.96 for July. The mean of the OMI–SSM/I difference is −0.58 cm in January and −0.41 cm in July, with a standard deviation of 0.47 cm and 0.45 cm, respectively.

AERONET is a network of globally distributed ground-based visible and near-IR sun photometers that measure atmospheric aerosol properties, inversion products, and precipitable water (aeronet.gsfc.nasa.gov) (Holben et al., 1998). Total water vapor column is retrieved from the 935 nm channel. The data used in this study are Version 2 daily averages. They are pre- and post-field calibrated, automatically cloud cleared and manually inspected.

Figure 11 shows the scatterplots of nearly coincident OMI and AERONET precipitable water for January and July in 2005 and 2006. All valid AERONET observations for the month are included for each panel. To find the nearly coincident observations, we use the OMI retrievals that have cloud fraction of <0.25 and are within a 0.5° radius of the AERONET site on the same day. There are typically 1–6 OMI data points for each AERONET data point, and they are averaged for comparison. Since artificial stripes in OMI swaths (Fig. 2) can significantly influence the comparison in this case, we perform an additional post-processing on OMI SCDs to remove the stripes by dividing a normalization vector. The normalization vector is derived using the mean of the middle third of the monthly averaged swaths and normalized so that the mean of the vector is unity. The de-striped SCDs are divided by the AMFs to convert to VCDs before subsequent processing. The correlation coefficients are in the range of 0.49 to 0.75. There appears to be better agreement in January than in July for both years. The slopes of the AERONET vs. OMI regression lines are <1. Considering that coincident OMI results have much larger uncertainty here than in Figs. 8–10 due to far fewer data points available...
for averaging, in addition to the different observational footprint and the highly variable nature of water vapor, the degree of agreement indicates that water vapor retrieval using OMI visible spectra is promising.

Figure 12 shows time-series comparisons between daily AERONET and OMI precipitable water for selected sites. This figure shows comparison not only of the mean but also of the day-to-day variation. The error bar for OMI in this plot only includes the uncertainty of the average of OMI SCDs. Other sources of error, including the error of AMF, the mismatch in timing between OMI and AERONET observation, the difference in observational footprint size, the spread due to scene inhomogeneity and the imperfection of the destripping procedure, are not included. Consequently, the total error for OMI should be larger than that shown in the figure. Despite this, we have found reasonably good matches between the two data sets. In the examples shown, the OMI result tracks both the mean and the variation of the AERONET result well except for occasional outliers. It is not surprising that we have also found examples where OMI does not agree with AERONET (not shown) due to the multiple error sources mentioned above. A comprehensive error analysis and data validation will be performed later.

5 Summary

Water vapor is an important molecule for weather, climate and atmospheric chemistry. There are distinct water vapor features in the OMI visible spectra that can be exploited to retrieve water vapor column amounts.

In this paper, we have presented our two-step operational OMI water vapor retrieval algorithm. We perform direct spectral fitting in the optimized spectral region of 430–480 nm to retrieve water vapor slant column density. This 50 nm long window includes the water vapor absorption feature at about 442 nm and 470 nm. Besides water vapor, we also fit O₃, O₂–O₂, NO₂, liquid water, the ring effect, the water ring effect and third-order closure polynomials. Our median retrieval uncertainty is about 1.2 × 10⁻²³ molecule cm⁻², about 50 % smaller than that obtained when using a shorter retrieval window. We have examined the sensitivity of our SCDs to the retrieval window, interfering molecules, reference spectra and other factors. Results show that it is important to include liquid water in our standard retrieval and use a relatively long retrieval window to reduce uncertainty. Results also show that the common mode over the ocean still has apparent structures as compared with that over the land, indicating the importance of improving the liquid water spectroscopy in this wavelength range.

We convert SCD to VCD by dividing by the AMF, which is a function of the scattering weight and shape factor. In our operational retrieval, we use a pre-calculated look-up table for the scattering weight and monthly mean assimilated water vapor profiles for the shape factor. We investigate the sensitivity of AMF to wavelength, solar zenith angle, surface albedo and cloud height. Results show that surface albedo and cloud information can lead to significant errors in AMF and therefore VCD. Our Level 2 product contains both scattering weights and AMFs in addition to VCDs for evaluation with and assimilation into models.

We compare our results with the MODIS near-IR data, GlobVapour combined MERIS+SSM/I product and AERONET measurements. Results show general agreement in terms of the spatial and temporal distribution both at the global level and for many sites. Future work will concentrate on further refining the retrieval algorithm, maintaining its long-term stability and performing extensive error analysis and data validation.

Acknowledgements. This study is supported by NASA's Atmospheric Composition: Aura Science Team program under grant NNX11A58G and the Smithsonian Institution. We acknowledge the OMI International Science Team for providing OMI data used in this study, the MODIS team for providing MODIS water vapor data, and the GlobVapour project for providing the combined MERIS+SSM/I water vapor product. We thank the AERONET project for its effort in establishing and maintaining measurement sites. We thank R. Lang and an anonymous reviewer for their reviews. We thank R. Lindstrot for providing comments.
H. Wang et al.: Water vapor retrieval from OMI visible spectra

Edited by: M. Weber

References


