Water vapor isotopologue retrievals from high-resolution GOSAT shortwave infrared spectra

C. Frankenberg¹, D. Wunch², G. Toon¹, C. Risi³, R. Scheepmaker⁴, J.-E. Lee¹, P. Wennberg², and J. Worden¹

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA
²California Institute of Technology, Pasadena, USA
³LMD/IPSL, CNRS, Paris, France
⁴SRON Netherlands Institute for Space Research, Utrecht, The Netherlands

Correspondence to: C. Frankenberg (christian.frankenberg@jpl.nasa.gov)

Received: 11 July 2012 – Published in Atmos. Meas. Tech. Discuss.: 6 September 2012
Revised: 19 December 2012 – Accepted: 21 December 2012 – Published: 7 February 2013

Abstract. Remote sensing of the isotopic composition of water vapor can provide valuable information on the hydrological cycle. Here, we demonstrate the feasibility of retrievals of the relative abundance of HDO (the HDO/H₂O ratio) from the Japanese GOSAT satellite. For this purpose, we use high spectral resolution nadir radiances around 6400 cm⁻¹ (1.56 μm) to retrieve vertical column amounts of H₂O and HDO. Retirements of H₂O correlate well with ECMWF (European Centre for Medium-Range Weather Forecasts) integrated profiles (r² = 0.96). Typical precision errors in the retrieved column-averaged deuterium depletion (δD) are 20–40 %. We compare δD against a TCCON (Total Carbon Column Observing Network) ground-based station in Lamont, Oklahoma. Using retrievals in very dry areas over Antarctica, we detect a small systematic offset in retrieved H₂O and HDO column amounts and take this into account for a bias correction of δD. Monthly averages of δD in the June 2009 to September 2011 time frame are well correlated with TCCON (r² = 0.79) and exhibit a slope of 0.98 (1.23 if not bias corrected). We also compare seasonal averages on the global scale with results from the SCIAMACHY instrument in the 2003–2005 time frame. Despite the lack of temporal overlap, seasonal averages in general agree well, with spatial correlations (r²) ranging from 0.62 in September through November to 0.83 in June through August. However, we observe higher variability in GOSAT δD, indicated by fitted slopes between 1.2 and 1.46. The discrepancies are likely related to differences in vertical sensitivities but warrant further validation of both GOSAT and SCIAMACHY and an extension of the validation dataset.

1 Introduction

Significant uncertainties in current climate model predictions of future warming are associated with the hydrological cycle, such as water vapor and cloud feedbacks (Held and Soden, 2006; Randall et al., 2007). For example, the variations in boundary layer cloud distribution are the largest source of spread in climate change projections (Bony and Dufresne, 2005; Bony et al., 2006). Atmospheric general circulation models (GCMs) must therefore accurately simulate the processes that control tropospheric humidity and clouds for their climate change predictions to be credible.

Measurements of stable isotopologues of water can provide key constraints on the processes controlling clouds and humidity, because (1) the isotopic composition of water will change as the water changes phase; (2) water sources such as the ocean and plants have different isotopic signatures (or compositions); and (3) mixing processes will affect the isotopic composition of water differently than phase changes. In the upper troposphere, the water vapor isotopic composition can be measured by satellites (Kuang et al., 2003; Nassar et al., 2007; Steinwagner et al., 2010) or in situ (Webster and Heymsfield, 2003; Sayres et al., 2010) and reflects the role of convection in the transport of water in the upper troposphere and through the tropopause (Moyer et al., 1996). In the mid-troposphere, the water vapor isotopic composition can also be measured by satellite (Worden et al., 2007; Herbin et al., 2009; Schneider and Hase, 2011) and in situ (Noone et al., 2011) and may give an indication about water cycle processes such as mixing of air masses (Galewsky and Hurley, 2010) or rain reevaporation (Worden et al., 2007).
In the lower troposphere, the water vapor isotopic composition is sensitive to the origin of water vapor and precipitation, such as continental recycling (Salati et al., 1979; Gat and Matsui, 1991) or air mass origin (Tian et al., 2001). To draw inferences about these lower tropospheric processes, many studies have so far relied on the isotopic composition of precipitation (e.g., Salati et al., 1979; Gat and Matsui, 1991; Tian et al., 2001). However, precipitation is strongly affected by post-condensation processes that blur the original vapor signal (Stewart, 1975; Lee et al., 2008; Risi et al., 2008). It is also spatially sparse and discontinuous in time as it relies on rainy days. The SCIAMACHY instrument onboard the European research satellite Envisat was the first instrument to provide global retrievals of water isotopes with high sensitivity near the ground (Frankenberg et al., 2009). Such retrievals open the possibility of exploiting the potential of water isotopes to better understand the lower tropospheric water budget, such as the role of continental recycling or convection in this budget (Risi et al., 2010). Owing to instrumental degradation, the current data record of SCIAMACHY HDO so far only covers the years 2003 through 2005 and connection to the entire Envisat satellite was irreparably lost in April 2012.

Here, we retrieve the column-averaged deuterium depletion of atmospheric water vapor using the Japanese Greenhouse gases Observing Satellite (GOSAT, Hamazaki et al., 2005; Kuze et al., 2009), which was launched on 23 January 2009 into a sun-synchronous orbit with a local overpass time of 13:00 LT. ≈10 000 soundings with 82 km² circular spatial footprints are recorded daily, repeating a regularly spaced global footprint grid every 3 days. Spectra are recorded by the TANSO Fourier Transform Spectrometer (FTS) onboard GOSAT. They enable independent retrievals of the total column amount of water vapor (H₂O) and its heavy isotope (HDO) in the 1.56 µm spectral region. This potential for GOSAT has so far not been exploited. We will discuss the retrieval setup in Sect. 2, validate and bias-correct the retrievals against ground-based observations in Sect. 3 and show global results and first comparisons with SCIAMACHY in Sect. 4.

2 HDO/H₂O retrieval setup from GOSAT

The column-averaged HDO abundance is often described in δ-notation, relative to standard mean ocean water (SMOW): δD = 1000 ‰ \left( \frac{VCD(\text{HDO})/VCD(\text{H}_2\text{O})}{R_i} - 1 \right), where VCD(HDO) and VCD(H₂O) denote the vertical column densities of HDO and H₂O, respectively. \( R_i = 3.1152 \times 10^{-4} \) is the ratio of HDO and H₂O for SMOW. VCD(H₂O) represents the vertically integrated total water vapor amount and is directly proportional to total precipitable water, a term more commonly used in hydrological sciences.

Our basic retrieval approach for δD largely follows retrievals from SCIAMACHY (e.g., Frankenberg et al., 2005a, 2008, 2009) using the IMAP-DOAS (Iterative Maximum a Posteriori–Differential Optical Absorption Spectroscopy) algorithm (Frankenberg et al., 2005b) with instrument-specific adaptions and ECMWF meteorological data for pressure, temperature and a priori humidity profiles. CO₂ is the only interfering species, and the disk-integrated solar line-list is adapted from the full-physics CO₂ retrievals used for GOSAT (originally developed for TCCON retrievals). The main difference with respect to the original IMAP-code is that we fit the spectra directly without taking the logarithm (however, as we use an iterative solver, the results are virtually identical). The continuum baseline is fitted with a second-order polynomial.

For the HDO a priori profile, we use a simplistic formulation, scaling the H₂O ECMWF prior profile with a pressure-dependent scaling factor, linearly varying from 0.9 at the surface to 0.5 at 50 hPa (i.e., a prior δD of -100 ‰ close to the surface and -500 ‰ in the stratosphere). In the state vector (both for HDO and H₂O), we include 10 retrieval layers, which are equidistant in pressure. The a priori covariance (1σ) was chosen to be 15 ‰ for all layers but the lowest, where a choice of 1500 ‰ as 1σ ensures that the retrieval is effectively unconstrained. HDO and H₂O are treated as independent species without any side constraints (besides the similarity in the prior profile) imposed on the HDO–H₂O functional relationship in the a priori covariance matrix. For the retrieval of δD, we use nadir radiance spectra measured in the shortwave infrared, viz. in the spectral range around 6400 cm⁻¹ (1.56 µm). The high spectral resolution of the TANSO FTS (2.5 cm maximum path difference resulting in 0.4 cm⁻¹ unapodized spectral resolution) enables the spectral fitting of weak HDO and H₂O lines in the aforementioned spectral range. Similar to retrievals of greenhouse gases in the near-infrared, the retrieval of profiles is not easily feasible due to low degrees of freedom and sensitivity to atmospheric scattering. Hence, we focus on retrieved total column amounts, which are more robust estimates in the near-infrared. An exemplary spectral fit using data near Lamont (Oklahoma) is shown in Fig. 1.

In the spectral fit, the weak CO₂ band around 6340 cm⁻¹ is included as it partially overlaps with HDO and H₂O lines. Even though spectral residuals (differences between modeled and measured radiances) are Gaussian and devoid of outliers, the reduced \( \chi^2 \) value is significantly higher than 1, mostly caused by low-frequency variations in the continuum (see upper panel). The Jacobians with respect to total column changes in HDO and H₂O can be seen in the lower panel. As for spectroscopic parameters, we use CO₂ absorption coefficient tables (including line-mixing effects) generated for the ACOS (Atmospheric CO₂ Observations from Space) and Orbiting Carbon Observatory (OCO-2) projects (Thompson et al., 2012). HDO and H₂O cross sections are calculated based on the line-list used for TCCON (Total Carbon Column Observing Network, Wunch et al., 2011a) using pure Voigt line-shapes. The TCCON H₂O and HDO linelist contains the Toth (2005) and Jenouvrier et al. (2007)
2.1 Quality filtering

We filter all retrievals based on a set of quality criteria, similar to O’Dell et al. (2012):

- \( \chi^2 < 3 \);
- standard deviation of residuum < \( 3 \times 10^{-9} \) W cm\(^{-1}\) sr\(^{-1}\) cm\(^{-1}\);
- relative error in retrieved HDO column < 15%;
- passed simple cloud filter (fitted O\(_2\) column amount > 90% of ECMWF O\(_2\) column amount);
- fitted H\(_2\)O column > 70% of ECMWF prior;
- 0.96 < CO\(_2\) ratio < 1.04;
- 0.8 < H\(_2\)O ratio < 1.2.

The CO\(_2\) and H\(_2\)O ratios are from IMAP-DOAS retrievals using the weak and strong CO\(_2\) bands as used in the full-physics ACOS retrievals (O’Dell et al., 2012; Crisp et al., 2012). These retrievals have been performed under a non-scattering assumption, and any strong deviation from this hypothesis leads to a divergence in retrieved abundances in the two bands. We found them to be particularly valuable in detecting strongly scattering scenes, including the detection of low-level clouds that can impact the retrieval accuracy. The simple cloud-filter based on the retrieved O\(_2\) column is identical to the one used in (Frankenberg et al. (2011) (see the Supplement)).

2.2 Bias correction

Systematic errors in spectroscopy and instrumental effects can bias both the H\(_2\)O and HDO vertical column densities. Errors in line-strengths typically result in constant multiplicative biases while errors in line-shape (e.g., pressure broadening) can create biases that depend on both viewing geometry (mostly air mass) and total H\(_2\)O column amount as these determine the degree of saturation in nadir-looking...
shortwave spectra. Constant multiplicative biases in the respective column densities will lead to a shift in $\delta D$ and are not too worrisome as the variability in $\delta D$ is unchanged and overall shifts are easy to calibrate out. Offsets in the retrieved column densities, however, can result in shifts that actually depend on the total amount of the retrieved columns. To investigate whether offsets are apparent in retrieved $\text{H}_2\text{O}$ and HDO column amounts, we extracted all land-only retrievals south of $60^\circ$S where the a priori ECMWF column amount was below $8 \times 10^{21}$ molecule $\text{cm}^{-2}$, i.e., a very dry atmosphere. This ensures a good fit of the intercept without a wet atmosphere dominating the slope fit, hence also the intercept. We then determined the offset by a least squares fit to a first-order polynomial, based on the ECMWF and retrieved column densities.

Results are shown in Fig. 2. We observe a significant negative offset for HDO and a small positive offset for $\text{H}_2\text{O}$. A bias-corrected $\delta D^\prime$ reads $\delta D^\prime = 1000 \%e \cdot \left( c_1 \cdot \text{VCD(HDO)} + 4.1 \times 10^{21} - 1 \right)$, with potential multiplicative errors $c_1$ in both vertical columns (caused by errors in line intensity and neither quantified nor applied in this study). The negative offset in HDO would, if uncorrected, result in unphysical $\delta D$ values and also lead to depletions approaching $-1000 \%e$ for column densities around $5 \times 10^{21}$ molecule $\text{cm}^{-2}$. In other words, the slope in a Rayleigh curve ($\delta D$ vs. $\text{H}_2\text{O}$) would be biased too steep, potentially affecting the interpretation of the results. We have neither found the root cause of the offset bias nor investigated the impact of erroneous broadening coefficients, which can lead to similar bias structures. Further complications may arise from the fact that both of these error terms can depend on the observed air mass, i.e., on season and latitude. However, the validation exercise in the following section will provide confidence that our simple bias correction yields very good agreement with ground-based observations over Lamont, Oklahoma.

### 2.3 Sensitivity to prior assumptions

We investigated the impact of a priori assumptions on the retrieved HDO and $\text{H}_2\text{O}$ quantities by comparing a limited GOSAT dataset around Lamont ($\approx 400$ soundings covering all seasons) using a widely different a priori profile assumption. We assumed a largely biased a priori profile with 1.5 times the ECMWF mixing ratios for $\text{H}_2\text{O}$ in the bottom half of the atmospheric column and an HDO profile that correspond to a $\approx 700 \%e$ depletion with respect to the ECMWF $\text{H}_2\text{O}$ profile at all layers.

Figure 3 shows that we are very insensitive to the a priori assumptions of the HDO and $\text{H}_2\text{O}$ profiles and that the standard deviation between a large set of retrievals using two widely different prior assumptions is only on the order of $10 \%e$. Boesch et al. (2012) also performed sensitivity studies with respect to the a priori HDO profile (but using a full multiple scattering code) and found similar ranges of the prior impact in most conditions (somewhat higher impact for a 5 % albedo case, but, for almost all terrestrial surfaces, the surface albedo in this spectral range is above 10 %). Undetected cirrus clouds can also bias the retrieval but mostly for very low surface albedos (5 %), which are unusual in the spectral range of interest.

### 2.4 Sensitivity to scattering processes and temperature

We refer the reader to Boesch et al. (2012) for investigating the impact of aerosols, temperature, pressure and cirrus clouds on the HDO retrievals. Uncertainties in the temperature profile appear to be a leading error source (presumably because high lower-state energies for water vapor transitions are considered here) but should be mostly a random component for larger averages (in time and space).

### 2.5 Other error sources

The primary source of error is single measurement noise, because $\text{H}_2\text{O}$ and in particular the HDO absorption lines are rather weak and signal-to-noise ratios of the recorded spectra are around 100–300. However, this error (typically around 20–100 $\%\text{e}$) is random and can be largely reduced by averaging in time and space. As for spectroscopy, we investigated the impact of using the official HITRAN database as well as the BXL database (Jenouvrier et al., 2007) and found a negligible impact on the retrieved variability in HDO/$\text{H}_2\text{O}$ abundances. Cross-correlations in retrieved HDO and $\text{H}_2\text{O}$ abundances are also negligible as the Jacobians for both are linearly independent (see Fig. 1). Systematic biases as observed over Antarctica, however, are hard to (a) detect and
For this proof of concept study, we chose a single TCCON station to validate GOSAT retrievals. For GOSAT satellite validation of greenhouse gases (e.g., Wunch et al., 2011b; Butz et al., 2011; Parker et al., 2011; Morino et al., 2011), Lamont had by far the largest number of coincidences, hence our choice to focus on this particular station. H₂O column amounts are validated against independent sonde measurements (Wunch et al., 2010), but HDO results from TCCON itself are so far unvalidated. Retrieval windows for all datasets used in this study are listed in Tables 1–2. For GOSAT and TCCON, there is a strong overlap of the retrieval windows in the 6300–6400 cm⁻¹ range.

We use the official TCCON GFIT algorithm to retrieve column amounts from high-resolution near-infrared spectra pointing directly at the sun’s center. The GFIT retrieval scales an a priori profile, preserving its original shape. The H₂O a priori profile is interpolated from NCEP 6-hourly data to the latitude, longitude, date, and local solar noon time at the site. The HDO a priori profile is generated from the NCEP H₂O profile using an empirically derived isotopic fractionation equation of HDO = H₂OVMR · 0.16 · (8.0 + log(H₂OVMR)). At the tropopause (H₂OVMR = 3 × 10⁻⁶), this equation gives HDO/H₂O = 0.4, whereas in the humid lower troposphere (H₂OVMR = 1 × 10⁻²) the equation gives HDO/H₂O = 0.96.

Owing to the overlap in retrieval windows, column averaging kernels for HDO are very similar for TCCON and GOSAT (see Fig. 4). Averaging kernels for H₂O should ideally have little impact on retrieved columns as the a priori profile (NCEP for TCCON and ECMWF for GOSAT) is relatively close to the truth as multiple sonde measurements per day over Lamont are assimilated into the numerical weather forecast fields. For TCCON and GOSAT H₂O retrievals, their respective averaging kernels are substantially different. For TCCON, the column averaging kernel resembles the one for HDO with increasing sensitivity at lower pressures. This is mainly related to the fact that very weak lines are fitted in a column scaling retrieval. For GOSAT, the H₂O kernels are close to unity throughout the column while the HDO kernel is very close to TCCON. It is important to note that the column averaging kernels are defined as changes in the retrieved total column with respect to a change in a sub-column at different altitudes. Under most conditions, the scale height of water vapor is only 1–2 km, and thus higher atmospheric layers (where the kernels can differ largely from 1) only contribute little to the total column and consequently have a small impact. Averaging kernels, however, are provided on a per sounding basis and thus allow users to take this effect into account.

Figure 5 shows a time series of retrieved HDO (scaled by $R_\lambda$) and H₂O column densities from GOSAT and the Lamont

---

**Fig. 3.** Sensitivity test for changes in a priori profile assumptions. The perturbed case scales the original H₂O prior profile with 1.5 above 500 hPa and assumes a constant $-700\%$ HDO depletion throughout the column. The standard deviation of the difference between the runs is approximately 10%.

**Fig. 4.** Typical TCCON and GOSAT column averaging kernels for H₂O and HDO.

**Table 1.** Retrieval windows for HDO and H₂O used for GOSAT and SCIAMACHY.

<table>
<thead>
<tr>
<th></th>
<th>GOSAT (cm⁻¹)</th>
<th>SCIAMACHY (cm⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6310–6440</td>
<td>4212–4248</td>
<td></td>
</tr>
</tbody>
</table>

---

(b) quantify. While we provide evidence regarding the usefulness of GOSAT data and its potential to reveal physical processes in the hydrological cycle, these caveats should be kept in mind and a more extensive validation exercise will have to be performed.
shows the time series of 

\[ \delta \text{HDO} \times 10^2 \]

shows the linear

\[ r^2 = 0.96, \text{slope}=0.89 \]

shows a scatter plot of GOSAT H vs. ECMWF H 

column averaged mixing ratio (ppm)

TCCON station with overlap from June 2009 through September 2011. Even though water vapor is highly variable and the coincidence criterion (±2‰) is not very stringent, the overall seasonal cycle shape and the inter-annual variability are consistent. The measurements show peaks in mid-summer and overall higher column amounts in 2010 than in 2009 or 2011.

Also the distinct summer peaks in 2010 and 2011 as well as the flat plateau for three consecutive months in 2009 are well reproduced. Values span a large dynamic range from about 0.1% column-averaged mixing ratio in winter to more than 0.6% in summer (individual retrievals can approach 1%).

Figure 6 shows a scatter plot of GOSAT vs. ECMWF H column amounts using all individual retrievals over Lamont (≈1400 soundings). The retrievals are well correlated (\( r^2 = 0.96 \)) albeit with a slope significantly smaller than unity, suggesting that H is low-biased by about 10% (resulting in a 100% high bias in δD caused by H2O; any additional errors in the HDO line-strength are difficult to quantify).

Figure 7 shows the time series of δD monthly means (and bias-corrected δD'). It is obvious that the raw retrievals (owing to the offset) lead to a depletion in winter that is too strong, while the bias-corrected δD exhibits similar magnitudes compared to TCCON. The overall shape of the seasonal cycle is extremely well reproduced in 2009 and early 2010, even showing unique month-to-month variability such as in October to November 2009 (with a sudden increase in δD even though the columnar water amount decreased).

In the second half of 2010 and early 2011, the correlation seems to decrease somewhat with GOSAT showing a distinct single minimum in the winter of 2010/2011. However, potential sampling biases (due to the large coincidence criteria) can also result in mismatches between the satellite- and the ground-based FTS. Also, both datasets exhibit an overall high bias as depletions around −50 to 0‰ in summer are unphysical over Lamont. Hence, we focus instead on variability rather than accuracy of absolute values.

Looking at the overall correlation between GOSAT and TCCON, we find a correlation coefficient (\( r^2 \)) of 0.79 and a slope of the linear fit of 1.15. Figure 8 shows the linear fit to the monthly averages. The raw retrievals are still very well correlated with TCCON but show a slope significantly larger than unity (i.e., too steep). The good correlation and slope close to unity provides confidence in the GOSAT retrievals and indicates that accurate water isotope retrievals in the shortwave infrared are possible using GOSAT. We acknowledge that HDO from TCCON has currently not yet
Fig. 7. Comparison of monthly δD over the Lamont (Oklahoma) TCCON station (GOSAT coincidences within ±2° latitude and longitude). Both means and medians are displayed, underlining the non-Gaussian nature of δ-D distributions in space and time.

validated itself against independent datasets and that more rigorous validation work, such as in Schneider et al. (2010), both on the TCCON and GOSAT side, will be necessary on top of more sensitivity studies related to the retrieval setup as well as choice of spectroscopic parameters.

4 Comparison with LMDZ model fields

We performed an initial comparison with the isotope-enabled LMDZ model (Risi et al., 2010) to provide further evidence that GOSAT data observe δD variations representing physical processes. Here, we look at typical Rayleigh curves for four distinct regions: tropical South America, Sahara, Southeast Asia and the continental US.

The top panel of Fig. 9 shows average Rayleigh curves and co-located model output for the four different regions. One can clearly see that the general behavior is very similar albeit a generally much steeper slope in GOSAT data. The amount effect in tropical regions (Southeast Asia and South America) can be clearly observed at very high columnar water vapor amounts. Also the typical differences between the regions are very similar. The bottom panel shows the deviations from the average Rayleigh curve of all four regions (restricted to vapor abundances observed in all regions).

![Graph of δD vs. H2O column abundance](image_url)

Fig. 8. Scatter plot and linear fit of monthly δD over the Lamont (Oklahoma) TCCON station (TCCON vs. GOSAT)

![Graph of Rayleigh curves](image_url)

Fig. 9. Average Rayleigh curves for land-only data of LMDZ and GOSAT over the Sahara (15°N–25°N, 5°W–20°E), continental US (30°N–50°N, 125°W–75°W), Southeast Asia (7°S–28°N, 70°E–120°E) and tropical South America (20°S–5°N, 120°W–40°W). The bottom panel shows the deviations from the average Rayleigh curve of all four regions (restricted to vapor abundances observed in all regions).
from an average Rayleigh curve are much more apparent. While work is still needed to investigate the cause of the differences in the slopes (be it a problem in the model, the measurement, or both), it provides confidence that GOSAT data reflect physical processes resulting in geographically dependent distributions of isotopic abundances.

5 Comparison with SCIAMACHY

SCIAMACHY is currently the only independent satellite dataset we can directly compare with as it also measures in the shortwave infrared (hence has a similar sensitivity). Unfortunately, currently available SCIAMACHY retrievals
do not coincide in time with GOSAT as detector degradation makes SCIAMACHY retrievals after 2005 complicated. An extension of the dataset is planned, but, so far, we can only compare the two datasets regarding the spatial distribution of $\delta D$ at seasonal timescales. For this particular comparison, we use recently improved SCIAMACHY retrievals from Scheepmaker et al. (2012).

Figure 10 shows global averages for December–February (DJF), March–May (MAM), June–August (JJA) and September–November (SON). For GOSAT, we relaxed the filter criteria (especially the $\text{CO}_2$ and $\text{H}_2\text{O}$ ratio, for which the maximal deviation from unity was changed to $< 0.0075$ and $< 0.5$, respectively) to achieve coverage in the tropics. Most regions, however, are not affected by this relaxation (e.g., the Lamont comparison is virtually unchanged by this choice). One can see that the GOSAT glint viewing geometry over subtropical oceans results in a much higher data yield. Over the oceans at higher latitudes, GOSAT takes regular nadir observations and retrievals are so far only possible over low cloud layers having passed the simple filter (similar to SCIAMACHY, which does not have a dedicated glint mode).

Overall, the spatial patterns as well as the seasonal variations are very similar, but a different value range for SCIAMACHY was chosen as $\delta D$ variability is somewhat lower in the SCIAMACHY retrievals. This manuscript is of a technical nature, but some aspects of the global distribution are worth discussing here: in boreal summer, highest $\delta D$ (larger than both over the oceans and tropical South America) is found in tropical Africa. This is most likely related to a higher degree of continental recycling over Africa (Brubaker et al., 1993), because transpiration is a non-fractionating process, which can result in lower atmospheric depletions than over oceans. Other aspects of the distribution and its application in studies of hydrological cycles are extensively discussed in, e.g., Berkelhammer et al. (2011); Lee et al. (2009); Risi et al. (2012b,a); Lee et al. (2012).

Also noteworthy from a more technical viewpoint is the fact that GOSAT retrievals are more reliable than SCIAMACHY over the oceans due to the glint viewing geometry. In addition, this is the case over land over some snow and ice areas such as Greenland, because both snow and ice have much lower surface albedos in the $2.3 \mu m$ range, making SCIAMACHY retrievals over these surface types problematic.

The lower variability of SCIAMACHY is apparent in Fig. 11: while correlation coefficients ($r^2$) of seasonal grid box averages vary between 0.65 in SON and 0.83 in JJA, the slopes vary between 1.20 and 1.46, showing that both datasets are strongly correlated but show a somewhat different variability in $\delta D$. Without the bias correction already applied to the GOSAT data, this discrepancy would increase by an additional 25%. One prime candidate for the discrepancies with SCIAMACHY is the averaging kernel for HDO for both GOSAT and TCCON. The column averaging kernels for HDO can well exceed unity while they are close to unity for $\delta D$. 

Fig. 11. Scatter plots of SCIAMACHY (2003–2005) and GOSAT (2009–2011) seasonal $2 \times 2$ degree grid box averages.
SCIAMACHY. This can amplify both the latitudinal gradients as well as the seasonal amplitude in both GOSAT and TCCON (which compare well with each other as averaging kernels are similar) but not in SCIAMACHY.

Despite the need for further research, the comparison with SCIAMACHY has shown that both datasets are well correlated even though they use entirely independent spectrometers and spectral regions. Apart from the aforementioned averaging kernels, spectroscopy, detector artifacts, differences in cloud filtering as well as spatial and temporal mismatches can all be causes for the current discrepancies. More validation using a consolidated set of ground-based observations will be necessary to rigorously evaluate and homogenize both data products. Also, model information on HDO profiles to estimate the effect of the column averaging kernels will be necessary to quantify this impact.

6 Conclusions

We have retrieved HDO and H$_2$O vertical column densities using shortwave infrared spectra from the Japanese GOSAT satellite. Column-averaged relative abundances of deuterium (δD) of atmospheric water vapor are retrieved with a precision that allows interpretation of single measurements (20–40‰). The enhanced precision, application of a different fitting window and small time overlap, is complementary to previous results from the SCIAMACHY instrument, which ceased operations in April 2012 and used the 2.3 μm range for HDO/H$_2$O retrievals. We compared GOSAT results against observations from the ground-based up-looking Fourier Transform Spectrometer site in Lamont (Oklahoma) and find a very good agreement in the timing and amplitude of the δD seasonal cycle. Seasonal averages on the global scale also compare very well with the SCIAMACHY instrument, albeit with higher variability observed with GOSAT (potentially caused by differences in vertical sensitivity). While more sensitivity studies and validation exercises are warranted, we show that global retrievals of the composition of water vapor isotopologues from GOSAT are feasible and can make valuable contributions to studies related to the hydrological cycle.

Appendix A

Derivation of effective pressure to include self-broadening effects

Under the assumption that the H$_2$O self-broadening coefficient is about 5 times larger than the air-broadening coefficient, the effective broadening coefficient in air (with H$_2$O) reads

\[ \gamma = \sum (\gamma_{n_i} \cdot \text{VMR}(n_i)) \]  

(A1)

\[ \approx \gamma_{\text{air}} \cdot (1 - \text{VMR}(\text{H}_2\text{O})) + \gamma_{\text{H}_2\text{O}} \cdot \text{VMR}(\text{H}_2\text{O}) \]  

(A2)

\[ = \gamma_{\text{air}} \cdot (1 - \text{VMR}(\text{H}_2\text{O})) + 5 \cdot \gamma_{\text{air}} \cdot \text{VMR}(\text{H}_2\text{O}) \]  

(A3)

\[ = \gamma_{\text{air}} \cdot (1 + 4 \cdot \text{VMR}(\text{H}_2\text{O})) \]  

(A4)

where $\gamma_{n_i}$ is the species ($n_i$) dependent broadening coefficient, $\gamma_{\text{air}}$ the air broadening coefficient and $\gamma_{\text{H}_2\text{O}}$ the water self-broadening coefficient. Since the actual width of the line is the product of $\gamma$ with air pressure, a multiplication of the pressure with $1 + 4 \cdot \text{VMR}(\text{H}_2\text{O})$ and subsequent use of just $\gamma_{\text{air}}$ leads to a good approximation of the final line-shape without the need to compute the self-broadening part explicitly.

Acknowledgements. Part of the research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. GOSAT Level 1B products (spectral data) were provided by the GOSAT Project (JAXA, NIES, and Ministry of the Environment Japan). US funding for TCCON comes from NASA’s Terrestrial Ecology Program (NNX08A186G), the Orbiting Carbon Observatory Program (NAS7-03001), the DOE/ARM Program and the Atmospheric CO$_2$ Observations from Space Program. R.S. was funded by the Netherlands Space Office as part of the User Support Programme Space Research under project GO-AO/16.

Edited by: M. Weber

References


Randall, D., Wood, R., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitma, A., Shukla, J., Srinivasan, J., Sumi, R. S. A.,


