Validation of two independent retrievals of SCIAMACHY water vapour columns using radiosonde data

A. du Piesanie¹, A. J. M. Piterš¹, I. Aben², H. Schrijver², P. Wang¹, and S. Noël³

¹Royal Netherlands Meteorological Institute (KNMI), de Bilt, The Netherlands
²Netherlands Institute for Space Research (SRON), Utrecht, The Netherlands
³Institute of Environmental Physics, University of Bremen, Bremen, Germany

Received: 26 November 2012 – Accepted: 10 December 2012 – Published: 21 January 2013
Correspondence to: A. du Piesanie (annelise.du.piesanie@knmi.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Two independently derived SCIAMACHY total water vapour column (WVC) products are compared with integrated water vapour data calculated from radiosonde measurements, and with each other. The two SCIAMACHY WVC products are retrieved with two different retrieval algorithms applied in the visible and short wave infrared wavelength regions respectively. The first SCIAMACHY WVC product used in the comparison is ESA’s level 2 version 5.01 WVC product derived with the Air Mass Corrected Differential Absorption Spectroscopy (AMC-DOAS) retrieval algorithm (SCIAMACHY-ESA). The second SCIAMACHY WVC product is derived using the Iterative Maximum Likelihood Method (IMLM) developed by Netherlands Institute for Space Research (SCIAMACHY-IMLM). Both SCIAMACHY WVC products are compared with collocated water vapour amounts determined from daily relative humidity radiosonde measurements obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) radiosonde network, over an 18 month and 2 yr period respectively.

Results indicate a good agreement between the WVC amounts of SCIAMACHY-ESA and the radiosonde, and a mean difference of 0.03 g cm$^{-2}$ is found for cloud free conditions. Overall the SCIAMACHY-ESA WVC amounts are smaller than the radiosonde WVC amounts, especially over oceans. For cloudy conditions the WVC bias has a clear dependence on the cloud top height and increases with increasing cloud top heights larger than approximately 2 km. A likely cause for this could be the different vertical profile shapes of water vapour and O$_2$ leading to different relative changes in their optical thickness, which makes the AMF correction method used in the algorithm less suitable for high clouds.

The SCIAMACHY-IMLM WVC amounts compare well to the radiosonde WVC amounts during cloud free conditions over land. A mean difference of 0.08 g cm$^{-2}$ is found which is consistent with previous results when comparing daily averaged SCIAMACHY-IMLM WVC amounts with ECMWF model data globally. Furthermore, we show that the measurements for cloudy conditions (cloud fraction $\geq 0.5$) with low
clouds (cloud pressure $\geq 930$ hPa) above the ocean and land compare quite well with radiosonde data.

1 Introduction

Water vapour is one of the most abundant constituents in the earth’s atmosphere. It is the most important greenhouse gas, because of its strong absorption of infrared radiation. The majority of water vapour is found in the lower parts of the atmosphere and its distribution is highly variable. Water vapour has a key influence on weather (clouds, precipitation) and atmospheric chemistry (e.g. the HO$_x$ cycle). Accurate measurements of water vapour with good spatial and temporal coverage are essential to monitor the distribution and variability of water vapour, and thereby the effect on climate, weather and chemistry.

SCIAMACHY (Scanning and Imaging Absorption Spectrometer for Atmospheric Chartography) onboard the European environmental satellite Envisat, uses atmospheric absorption spectra to derive vertical total water vapour column (WVC) amounts. In this study two independently derived SCIAMACHY total WVC products retrieved with two different retrieval algorithms applied in the visible and SWIR wavelength regions respectively and each with their own error characteristics, are compared to radiosondes and to each other.

The SCIAMACHY product OL version 5.01, distributed by ESA, includes a total water vapour column retrieved with the AMC-DOAS (Air Mass Corrected Differential Absorption Spectroscopy) method. The AMC-DOAS retrieval algorithm, developed by the University of Bremen, was first used to derive water vapour column amounts from GOME (Global Ozone Monitoring Experiment) onboard ERS-2 (Noël et al., 1999) and later SCIAMACHY (Noël et al., 2004, 2005) and GOME-2 on MetOp (Noël et al., 2008). Mieruch et al. (2010) found a systematic difference between water vapour derived from GOME/SCIAMACHY with AMC-DOAS v1.0 and water vapour derived from SSM/I (Special Sensor Microwave/Imager) with the HOAPS (Hamburg Ocean Atmosphere
Parameters and Fluxes from Satellite) v3.1 algorithm. This difference was on average 1–2 kg m\(^{-2}\) for clear-sky situations and up to 8–10 kg m\(^{-2}\) along the ITCZ (Intertropical Convergence Zone) and the SPCZ (South Pacific Convergence Zone) and in tropical regions of the China Sea, the AMC-DOAS values being lower. The major reason for this difference was believed to be that the correction for partly cloudy scenes, used in the AMC-DOAS method, was not large enough.

SCIAMACHY water vapour columns are also derived by the Netherlands Institute for Space Research (SRON) with the Iterative Maximum Likelihood Method (IMLM). This algorithm is applied to the near-infrared wavelength region and has been successfully used to retrieve CO (Gloedemans et al., 2008). More recently it has been used to retrieve H\(_2\)O (Shrijver et al., 2009) using the spectral range between 2353 nm and 2368 nm. Schrijver et al. (2009) compared the IMLM water vapour columns with ECMWF data and found an average difference of 0.1 kg m\(^{-2}\) for IMLM water vapour columns in clear-sky situations over land.

In this paper SCIAMACHY total WVC measurements are compared with integrated water vapour data obtained from collocated radiosonde measurements respectively covering an 18 month and 2 yr period.

In Sect. 2 the data sets used are described in more detail. The comparisons to radiosondes are performed in Sects. 3 and 4, and in Sect. 5 both datasets are compared to each other.

## 2 Data sources

### 2.1 Radiosondes

The two individual SCIAMACHY water vapour data sets are compared with integrated water vapour derived from radiosonde relative humidity measurements. The radiosonde measurement data used here were acquired from the European Centre for Medium-Range Weather Forecasts (ECMWF) radiosonde network (data obtained
from ECMWF’s Ecgate server https://ecaccess.ecmwf.int, registered users only). This network consists of daily radiosonde measurements from a large number of globally distributed ground measurement sites (Fig. 1). The number of globally performed radiosonde measurements available on a daily basis varies and not all measurement sites necessarily perform a radiosonde measurement every day, or alternatively some sites perform more than one measurement per day for example during measurement campaigns. In general balloon-borne radiosondes are launched twice daily at 12:00 UTC and 00:00 UTC and depending on the size of the balloon they can reach altitudes up to 30 km. These global radiosonde humidity measurements are collected by various sonde types with different sensor characteristics. A number of methods have been developed to correct for biases in humidity observations, however many of these methods focus primarily on individual radiosonde instruments (Wang and Zhang, 2008). Overall the majority of radiosonde instrument types show a dry bias, predominantly in the upper troposphere where the bias can reach 5% to 8% in relative humidity (Sun et al., 2010). Each relative humidity profile obtained from the radiosonde measurements is integrated to a vertical column amount, provided the profile has enough measurement layers. The integration is performed using the pressure and specific humidity at each level in the vertical. Specific humidity is calculated using the relative humidity, pressure and saturated vapour pressure, which in turn is calculated according to the Goff-Gratch equation (Goff and Gratch, 1946).

2.2 SCIAMACHY water vapour columns

The SCIAMACHY instrument on board the European Space Agency’s (ESA) Envisat satellite was launched in 2002 and stopped operations April 2012. SCIAMACHY measures Earthshine radiances and solar irradiance spectra from the ultraviolet, visible and near-infrared wavelength region (240–2380 nm) in limb and nadir viewing geometry (Bovensmann et al., 1999; Gottwald and Bovensmann, 2011). Envisat is operated in a sun-synchronous orbit and has an equator crossing time of 10:00 LT. Individual ground pixels size depends on the selected integration time (Gottwald and...
Bovensmann, 2011). In the wavelength range between 2353 to 2368 nm where the IMLM is applied to derive the total WVC, a SCIAMACHY ground pixel typically has a spatial resolution of 120 km × 30 km (Schrijver et al., 2009). In the shorter wavelength range where the AMC-DOAS method is applied, a ground pixel has a typical size of 60 km × 30 km (Noël et al., 2004).

2.2.1 ESA product, version 5.01, AMC-DOAS method

Total WVC amounts in the current SCIAMACHY level 2 version 5.01 ESA product are derived with the AMC-DOAS retrieval algorithm version 1.0 as developed by Noël et al. (2004). The AMC-DOAS method is a modified approach of the Differential Optical Absorption Spectroscopy (DOAS) method, using the differential absorption structures to derive total columns. The method is applied in a spectral fitting window between 688 nm and 700 nm. Both O₂ and water vapour absorb in this spectral region and have similar optical depths. The method is described by the following equation:

\[
\ln \left( \frac{I}{I_0} \right) = P - a \left( \tau_{O_2} + cC_v^b \right)
\]

where \( I \) and \( I_0 \) represent the measured earthshine radiance and irradiance spectra respectively, \( \tau_{O_2} \) represents the O₂ optical depth, \( C_v \) denotes the vertical water vapour column, \( c \) is the water vapour air mass factor, \( b \) is a parameter correcting for saturation effects in the water vapour absorption lines, and the scalar factor \( a \) represents the so-called air mass factor correction factor (AMF CF). All spectral broadband contributions resulting from Rayleigh and Mie scattering or surface albedo are approximated by a polynomial \( P \). The parameters \( b, c, \) and \( \tau_{O_2} \) are calculated from radiative transfer calculations, and \( P, a, \) and \( C_v \) are fitted. The value of the AMF CF \( a \) is mainly determined by the O₂ absorption features in the fitting window. It serves as a first order correction factor for variations in, e.g. cloud cover and surface albedo, with respect to the model atmosphere, assumed within the radiative transfer calculations, i.e. a cloud-free tropical background atmosphere, a surface albedo of 0.05 and a surface elevation of 0 km.
This correction is expected to work well when the air mass factors for O\textsubscript{2} and water vapour are affected by such variations in a similar way.

The SCIAMACHY-ESA total WVC product excludes measurements for which the AMF CF is smaller than 0.8 or where the solar zenith angle is larger than 88°. For this study SCIAMACHY total water vapour column data covering the 18 month period from February 2010 to mid-August 2011 is used.

2.2.2 SRON product, version 7.4.1, Iterative Maximum Likelihood Method

The Iterative Maximum Likelihood Method (IMLM) developed by Netherlands Institute for Space Research (SRON) for retrieving trace gas columns from the near-infrared is used to derive total water vapour columns (Schrijver et al., 2009). The IMLM is applied in the wavelength range between 2353 nm and 2368 nm and because of the overlapping H\textsubscript{2}O, CO and CH\textsubscript{4} absorption lines in this region all three species are retrieved simultaneously. The IML method (version 6.3) is described in detail in Gloudemans et al. (2008). It is based on scaling a priori atmospheric profiles, and a model of the expected detector signal is fitted to the measurements by adjusting the total column amounts of the trace gases (H\textsubscript{2}O, CO and CH\textsubscript{4}) that play a role in this particular retrieval window (Schrijver and Gloudemans, 2008; Schrijver et al., 2009). Water vapour and temperature profiles used in the scaling procedure are from collocated ECMWF analysis. All subsequent updates performed in version 7.4 are described in detail in Gloudemans et al. (2009). One of these updates is the use of a different spectroscopic database of H\textsubscript{2}O and CH\textsubscript{4} for calculating the cross-sections and line broadening (Gloudemans et al., 2009). Schrijver et al. (2009) found that using this updated spectroscopic parameters for H\textsubscript{2}O leads to an improvement in the IMLM WVC product when comparing it with ECMWF model data.

The retrieval has no cloud correction, so it is expected that errors generally are smaller for cloud-free conditions than for (partly) cloudy conditions, when a major part of the water vapour column is hidden below the clouds. In case of cloud-free conditions, the signal-to-noise ratio of the measurement is closely related to the surface
albedo. This leads to smaller errors for measurements taken over land than over ocean (Schrijver et al., 2009).

Therefore the first comparison in Sect. 4 is performed for cloud-free pixels over land only. The cloud filter included in the product is based on the SCIAMACHY polarisation measurement device (PMD) Identification of Clouds and Ice/snow (SPICI) cloud filter as described by Krijger et al. (2005). This method distinguishes between cloud free and partly cloudy scenes (de Laat et al., 2007). An additional check on the cloud filter is performed by comparing the simultaneously retrieved methane columns with the expected methane columns based on the ECMWF surface pressure. Cases where the retrieved methane measurements are more than 10% below the expected value, taking into account surface pressure and elevation, are not included. The usefulness of this additional methane filter has been shown by Schrijver et al. (2009). For the comparisons with radiosondes the two years of 2004 and 2009 of SCIAMACHY-IMLM total WVC data are used.

3 Comparison of SCIAMACHY-ESA WVC with radiosondes

For the comparison of water vapour column measurements between SCIAMACHY-ESA WVC data and the radiosonde data, the following collocation criteria were used: a spatial difference of less than 100 km (as calculated from the centre of a SCIAMACHY pixel) and a time difference of less than 3 h. Using these criteria a total of 50 470 individual measurements were found. A time series of the global daily mean difference between the two data sets and standard deviations are shown in Fig. 2. Gaps in the time series indicate missing data (in total 84 days) from either the SCIAMACHY or radiosonde water vapour data sets. The SCIAMACHY-ESA WVC amounts are overall smaller than those of the radiosondes, and there seems to be a small seasonal variation, with a slightly larger bias during July–September and slightly smaller during March–May. A generally good agreement (correlation coefficient $r = 0.89$) between individual collocated cases of the data sets can be seen from Fig. 3 (left) and a mean
difference of $-0.32 \pm 0.01$ g cm$^{-2}$ is found, consistent with values reported by Mieruch et al. (2010), when comparing the SCIAMACHY AMC-DOAS version 1.0 from the University of Bremen to SSM/I data. The scatter of the data is 0.69 g cm$^{-2}$, slightly larger than previous comparisons performed between AMC-DOAS and ECMWF model water vapour data (Noël et al., 2005), but it should be noted that the scatter reported by Noël et al. (2005) was derived for global daily means instead of individual collocations as done here. The scatter is slightly less (standard deviation = 0.6 g cm$^{-2}$) when considering only cases where the cloud fraction is equal to zero (10022 individual collocations, Fig. 3 right). In this case a mean difference of $0.03 \pm 0.01$ g cm$^{-2}$ is found. The cloud parameters used in this section are derived from the SCIAMACHY operational products. Cloud coverage is retrieved using the Optical Cloud Recognition Algorithm (OCRA) (Loyola, 1998) and Semi-Analytical Cloud Retrieval Algorithm (SACURA) (Kokhanovsky et al., 2005) is used to derive cloud optical thickness and cloud top height.

To identify what could be responsible for the differences in WVC amounts as observed in Fig. 3, the bias as a function of several parameters, including various cloud properties, is analysed. Investigating the bias as a function of solar zenith angle, albedo, cloud optical depth and cloud fraction (other than selecting for cloud free cases only as done above) does not show a strong relation between these parameters and the WVC differences (Fig. 4). This suggests that the AMF CF on average works well for variations in the parameters such as solar zenith angle with respect to the assumptions in the radiative transfer calculations (see Sect. 2.2.1). A strong dependence is found when investigating the bias as a function of cloud top height during cloudy conditions. In Fig. 5 this relation is plotted for all points with a cloud cover fraction more than or equal to 0.9. For cloud top heights larger than approximately 2 km the bias increases rapidly with increasing cloud top heights. A likely explanation for this increasing bias with increasing cloud top height during very cloudy conditions can be found in the very different vertical profile shapes of O$_2$ and water vapour. O$_2$ is well-mixed in the atmosphere, while water vapour rapidly decreases with altitude. If the presence of a cloud,
the cloud shields the water vapour and $O_2$ below it and the optical thickness for both decreases. Due to the differences in profile shape, the optical thickness of water vapour decreases more than the optical thickness of $O_2$, and this will lead to an AMF CF which is too large (see Eq. 1), and a retrieved water vapour column which is too small. This effect is illustrated in Fig. 6, which shows an estimate of the error in the water vapour column as a function of cloud top height, as calculated with the Doubling-Adding KNMI (DAK) radiative transfer model (de Haan et al., 1987; Stammes, 2001).

A relationship between the bias in water vapour column and cloud properties was also reported by Mieruch et al. (2010). They found that the average bias between GOME/SCIAMACHY water vapour and the HOAPS product over ocean depends on monthly mean cloud water path, for cases where the cloud water path was smaller than 30 g·m$^{-2}$ (i.e. for 12% of the data). The largest biases were found over the ITCZ, the SPCZ, and tropical regions of the China Sea. Mieruch et al. (2010) concluded that this bias is caused by the AMF CF, which does not adequately compensate for the occurrence of clouds. With the results found here, we can slightly refine this conclusion: for clouds below 2 km, the AMF CF actually performs quite well, and the large bias found over the ITCZ, the SPCZ, and the tropical regions of the China Sea are probably due to relatively high cloud altitudes, as can be seen from Fig. 7. This figure also shows that there is a relation between average cloud top pressure and cloud water path, which explains the dependence of the bias on cloud water path found by Mieruch et al. (2010).

Figure 8 shows the average difference between SCIAMACHY-ESA and radiosonde WVC as a function of the AMF CF. For values above 1.0, the average difference does not significantly differ from 0, suggesting that in these situations the behaviour of the optical thickness of water vapour and oxygen are somewhat similar. However, for smaller values of the AMF CF, which typically occur in the presence of (high) clouds, the AMF CF should in fact be even smaller to account for the difference in profile shapes of water vapour and oxygen, as argued above. Table 1 summarises the results when applying various selection criteria based on the cloud fraction, cloud height and the AMF CF.
Examining the WVC measurements over land and ocean shows a noticeable disparity between the WVC differences over these surfaces. The mean WVC differences shown in Fig. 9 are for each of the collocated radiosonde stations for the 18-month period for cloud free conditions (according SCIAMACHY operational cloud products). It should be noted that radiosonde measurements over the ocean are limited, and many stations are close to coastal areas. The SCIAMACHY topography is used to distinguish between the land and ocean SCIAMACHY-EPA measurements. Radiosonde stations are designated as land (or ocean) stations if more than 90% of the collocated SCIAMACHY-EPA measurements occur over land (or ocean). In general mostly negative values are found over the ocean and positive values over land. The weighted mean WVC difference amounts to 0.11 g cm⁻² over land and -0.49 g cm⁻² over ocean, where the number of measurements per radiosonde station has been used as a weight. The SCIAMACHY-EPA WVC measurements are further compared to WVC data from ECMWF over land and ocean surfaces. The SCIAMACHY-EPA WVC amounts are compared with ECMWF data for a one month period for cloud free conditions (according to SPICI cloud filter, see Sect. 2.2.2) where both data sets have been averaged to a ground pixel size ranging from 120 km x 30 km to 240 km x 30 km (at higher latitudes) (Fig 10). The result similarly indicates a division between land and oceans, with positive bias over land and negative bias over the ocean ranging between values of 1 g cm⁻² over land to -1 g cm⁻² over oceans.

4 Comparison of SCIAMACHY-IMLM WVC with radiosondes

The same distance and time collocation criteria as described in Sect. 3 and used for comparing ESA WVC amounts with those from radiosondes, are used here for the SCIAMACHY-IMLM WVC amounts. A number of selection criteria as suggested by Schrijver et al. (2009) are applied to the collocated measurement data. This includes filtering the data to only include cloud free cases by selecting data according to the SPICI cloud product and including the additional criterion of the simultaneously retrieved
methane total column (described in Sect. 2.2.2), excluding data measurements performed over the ocean as well as measurements with an average signal strength below 100 binary units per detector pixel (BUpx\(^{-1}\)). When filtering the collocated data purely according to its quality flag and excluding cases below a signal level of 100 BUpx\(^{-1}\), a total of 48,844 cases are found for the two year period (2004 and 2009). When further selecting from this data by only including cases over land and cloud-free conditions according to the SPICI cloud product, a total number of 6,590 cases remain. For this selection a relatively small mean difference of 0.01 (± 0.007) g cm\(^{-2}\) is found between the SCIAMACHY-IMLM and the radiosonde WVC amounts, but with a high standard deviation of 0.56 g cm\(^{-2}\). Finally applying the additional methane column criteria, as was done in the study of Schrijver et al. (2009), only 3,489 collocations are found. For this selection, the individual cases compare well (Fig. 11) and a mean difference of 0.08 (± 0.007) g cm\(^{-2}\) is found, consistent with the value found by Schrijver et al. (2009) when comparing daily averaged SCIAMACHY-IMLM WVC amounts with ECMWF model data globally. The standard deviation is 0.42 g cm\(^{-2}\), which is slightly larger than that found by Schrijver et al. (2009; generally below 0.3 g cm\(^{-2}\)).

The SCIAMACHY-IMLM water vapour product is expected to have the smallest measurement noise error when the measured signal strength is relatively large. In principle, the presence of clouds will lead to larger signals and thus to smaller measurement noise errors. But clouds also shield part of the water vapour column, so that the most accurate total water vapour column retrieval is expected under cloud-free conditions and areas with bright surfaces. The SCIAMACHY-IMLM WVC measurement noise error shows a clear dependence on the signal strength (Fig. 12), indicating larger measurement noise errors for cases with lower signal strengths. The increase in the measurement noise error from the year 2004 to 2009 can be attributed to detector degradation caused by radiation damage to the individual detector pixels (Kleipool et al., 2007). Gloudemans et al. (2008) showed a dependence of the measurement noise error on the surface albedo for CO and CH\(_4\). For cases with low albedo values a large measurement noise error is thus expected, for example measurements taken over cloud
free oceans. The mean bias and the standard deviation between the WVC amounts of SCIAMACHY-IMLM and that of radiosondes are related to the SCIAMACHY-IMLM signal strength, as can be seen in Fig. 13. The black lines in the figure indicate cases chosen according to the SPICI cloud free criterion. The red lines are for cases where FRESCO+ (Fast Retrieval Scheme for Clouds from the Oxygen A-band) (Wang et al., 2008) cloud fraction is smaller than 0.02. FRESCO+ data were obtained from TEMIS (www.temis.nl). For this selection only 4% of the points have signal strengths larger than 1000 BUpx$^{-1}$, compared to 21% of the selection based on the SPICI cloud free criterion. The largest signal strengths most likely indicate the presence of some small or scattered cloud still present in the selection. The FRESCO+ cloud parameters are averaged over the ground pixels to match the larger SCIAMACHY-IMLM water vapour ground pixel. For cloud-free conditions (FRESCO+ cloud fraction smaller than 0.02) SCIAMACHY-IMLM WVC are consistent with those integrated from radiosonde WVC amounts, with a random uncertainty depending on the SCIAMACHY signal strength, both for pixels over land and over ocean. Also this systematic bias seems to depend on signal strength. Using the cloud parameters from FRESCO+ as opposed to that of SPICI leads to less scatter and a slightly smaller mean bias.

The suggested selection criteria whereby cases with signal strength below 100 BUpx$^{-1}$ are excluded eliminate the majority of measurement cases over ocean surfaces (due to low surface albedo over ocean surfaces leading to a low signal to noise ratio and therefore larger errors). Similarly, excluding measurements taken during cloudy conditions also largely reduces the number of useable data. Scenes with sufficient cloud have a higher albedo than that of ocean surfaces (typically 0.3 to 0.5 as opposed to smaller than 0.01) leading to larger signals and thus smaller measurement noise errors. Therefore including certain cloud cases over land and ocean surfaces in the data selection could be advantageous. Gloudemans et al. (2009) showed that CO columns over the ocean for cloudy conditions can be retrieved successfully by using the cloud top heights calculated from the simultaneously retrieved CH$_4$ column. As water vapour has a large mixing ratio close to the surface and decreases rapidly...
with altitude, only scenes with low clouds are included in this particular comparison. Figure 14 shows the comparison between SCIAMACHY-IMLM and radiosonde WVC amounts for land and ocean cases respectively. Measurements are filtered to include cases with FRESCO+ cloud height pressures larger than 930 hPa and cloud cover fraction larger than 0.5. The WVC amounts compare relatively well for both land and ocean comparisons, with a bias of $-0.44 \pm 0.05 \text{ g cm}^{-2}$ for land cases and $-0.27 \pm 0.04 \text{ g cm}^{-2}$ for ocean cases, but with some scatter. The larger bias and scatter over land might be due to the larger variability of water vapour below the clouds, as compared to over ocean. Again it should be noted that radiosonde measurements over the ocean are limited, and many collocations are close to coastal areas. To characterise the uncertainty of water-vapour above the cloud, the radiosonde obtained relative humidity profile is integrated from the top of the cloud rather than from the ground surface as done with previous comparisons. The top of the cloud is determined by the FRESCO+ cloud pressure. These results (Fig. 15) show, when comparing the partial WVC from radiosondes to SCIAMACHY-IMLM WVC for cloudy conditions with various cloud heights, a mean bias of $-0.09 \text{ g cm}^{-2}$ for cases over land and $-0.06 \text{ g cm}^{-2}$ for cases over ocean are found. In general the WVC amounts obtained from the partially integrated radiosonde profile are higher than that of SCIAMACHY-IMLM. Wang et al. (2008) has found that cloud height retrieved by FRESCO+ is in reality closer to the middle of the cloud than the top of the cloud. This could be a possible explanation as to the larger radiosonde WVC amounts that are found when comparing it to SCIAMACHY-IMLM WVC amounts above cloud. The selections used in the above two comparison cases are summarised in Table 2 and show that it can be beneficial to extend the SCIAMACHY-IMLM water vapour data set by including SCIAMACHY-IMLM WVC measurements taken during certain cloudy conditions above ocean and land surfaces. This can be done by either selecting cloudy conditions with low cloud height or alternatively making use of a partially integrated radiosonde water vapour profile.
5 Inter-comparison between SCIAMACHY AMC-DOAS and IMLM water vapour data

A comparison between the retrieved SCIAMACHY-IMLM and SCIAMACHY-ESA water vapour column data is performed for a one day period for 6 June 2010. In general the pattern of water vapour from both products on this day seems reasonable (Fig. 16), with higher water vapour amounts in the tropics and low amounts at higher latitudes. Figure 17 shows the comparison of SCIAMACHY-ESA WVC with SCIAMACHY-IMLM WVC amounts for the one day period. Here the SCIAMACHY-ESA WVC amounts are averaged over the ground pixels to match the larger SCIAMACHY-IMLM water vapour ground pixel. A total of 1449 collocated cases are found for ground pixels over land with SCIAMACHY-IMLM signal level larger than 100 BU px$^{-1}$ for cloud free conditions according to the SPICI cloud product and using the additional methane criterion (described in Sect. 2.2.2). Based on this selection the SCIAMACHY-ESA WVC amounts are generally slightly higher than that of SCIAMACHY-IMLM amounts, however the correlation between the two data sets is rather high ($r = 0.96$).

6 Conclusions

In this study the two independently derived SCIAMACHY total WVC products from two different spectral ranges are compared with integrated water vapour data obtained from collocated radiosonde relative humidity measurements, and with each other. In general a good agreement is found when comparing individual collocated radiosonde WVC to that of SCIAMACHY-ESA WVC amounts and a mean difference (SCIAMACHY-ESA minus radiosonde WVC) of $-0.32$ g cm$^{-2}$ is found. The scatter of the data is $0.69$ g cm$^{-2}$, slightly larger than previous comparisons between globally averaged daily means ECMWF model water vapour data and AMC-DOAS (Noël et al., 2005). The mean bias and scatter improves ($0.03$ g cm$^{-2}$ and $0.6$ g cm$^{-2}$ respectively) when considering only cases with cloud free conditions. The bias has a clear
dependence on the cloud top height and increases with increasing cloud top heights larger than approximately 2 km. A likely explanation for the bias is the very different vertical profile shapes of O$_2$ and water vapour, ultimately leading to an AMF CF which is too large. In order to minimise the bias it could be advantageous to filter SCIAMACHY-ESA water vapour data according to the cloud top height or the AMF CF (for example 0.9 or 0.95). Furthermore, a division between the WVC differences over land and ocean are found which is also confirmed when comparing SCIAMACHY-ESA with ECMWF WVC amounts in June 2005, where a positive bias is observed over land and a negative bias is observed over the ocean.

Applying the filtering criteria as suggested in Schrijver et al. (2009) to the SCIAMACHY-IMLM WVC data largely decreases the amount of useable data, however a good agreement is found when comparing these individual WVC amounts to that of radiosonde WVC. A mean difference of 0.08 g cm$^{-2}$ is found, consistent with the value found by Schrijver et al. (2009), but with a slightly higher standard deviation of 0.42 g cm$^{-2}$ than that found by Schrijver et al. (2009). The strict filtering criteria exclude all measurements above ocean surfaces. In Sect. 4 we show that the measurements for low clouds ($\geq 930$ hPa) above the ocean actually compare quite well with radiosonde data, albeit with a large negative bias and scatter, probably caused by the invisible water vapour below the cloud. A comparison of SCIAMACHY-IMLM WVC amounts measured above clouds also show a good agreement when comparing it to the WVC obtained by integrating the radiosonde profile from the cloud top pressure (instead of the ground surface). The small bias is most likely caused by the difference between the retrieved cloud top height and effective scattering height within the cloud.

SCIAMACHY-ESA WVC amounts are generally higher than that of SCIAMACHY-IMLM amounts when comparing the two WVC products with each other, with a good correlation found between the two WVC data sets.
Acknowledgements. SCIAMACHY is a joint project of the German Space Agency DLR and the Netherlands Space Office NSO with contribution from the Belgian Space Agency. The work performed is partly financed by Netherlands Space Office via the SCIAVISIE Science project.

References


Validation of SCIAMEHY water vapour columns

A. du Piesanie et al.


Table 1. Applying various selection criteria on the SCIAMACHY-ESA and radiosonde WVC data sets. The number of cases in selection, mean WVC difference, standard deviation and correlation coefficient are given in each column.

<table>
<thead>
<tr>
<th>Selection criteria</th>
<th>No. of cases in selection</th>
<th>WVC mean difference (g cm(^{-2}))</th>
<th>Standard deviation</th>
<th>Correlation coefficient (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All individual cases</td>
<td>50 470</td>
<td>−0.32</td>
<td>0.69</td>
<td>0.89</td>
</tr>
<tr>
<td>Cloud height ≤ 2 km or cloud fraction ≤ 0.2</td>
<td>33 518</td>
<td>−0.13</td>
<td>0.57</td>
<td>0.92</td>
</tr>
<tr>
<td>AMF CF ≥ 0.95</td>
<td>33 107</td>
<td>−0.09</td>
<td>0.48</td>
<td>0.92</td>
</tr>
<tr>
<td>Cloud height ≤ 2 km</td>
<td>24 924</td>
<td>−0.05</td>
<td>0.5</td>
<td>0.93</td>
</tr>
<tr>
<td>Cloud fraction = 0</td>
<td>10 022</td>
<td>0.03</td>
<td>0.6</td>
<td>0.89</td>
</tr>
</tbody>
</table>
Table 2. Applying various selection criteria on the SCIAMACHY-IMLM and radiosonde WVC data sets. The number of cases in selection, mean WVC difference, standard deviation and correlation coefficient are given in each column.

<table>
<thead>
<tr>
<th>Selection criteria</th>
<th>No. of cases in selection</th>
<th>WVC mean difference (g cm(^{-2}))</th>
<th>Standard deviation</th>
<th>Correlation coefficient (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection criteria as used by Schrijver et al. (2009)(^a)</td>
<td>3489</td>
<td>0.08</td>
<td>0.42</td>
<td>0.9</td>
</tr>
<tr>
<td>CLpress ≥ 930 hPa, CLfrac ≥ 0.5 over land(^b)</td>
<td>337</td>
<td>−0.44</td>
<td>0.86</td>
<td>0.8</td>
</tr>
<tr>
<td>CLpress ≥ 930 hPa, CLfrac ≥ 0.5 over ocean(^b)</td>
<td>110</td>
<td>−0.27</td>
<td>0.44</td>
<td>0.92</td>
</tr>
<tr>
<td>Partial WVC, CLpress ≥ 930 hPa, CLfrac ≥ 0.5 over land(^b)</td>
<td>9789</td>
<td>−0.09</td>
<td>0.34</td>
<td>0.83</td>
</tr>
<tr>
<td>Partial WVC, CLpress ≥ 930 hPa, CLfrac ≥ 0.5 over ocean(^b)</td>
<td>1436</td>
<td>−0.06</td>
<td>0.34</td>
<td>0.83</td>
</tr>
</tbody>
</table>

\(^a\) SCIAMACHY-IMLM data selected to include measurements over land and cloud free conditions according to SPICI cloud product, applying the additional methane column criterion and excluding cases with a signal level smaller than 100 BUpx\(^{-1}\).

\(^b\) CLfrac, Cloud fraction; CLpress, Cloud pressure.
Fig. 1. Location of the radiosonde stations used in both the SCIAMACHY-ESA and SCIAMACHY-IMLM water vapour column inter-comparisons with radiosonde WVC data.
Fig. 2. The global daily difference between SCIAMACHY-ESA WVC amounts and radiosonde WVC as a function of time (1 February 2010 to 10 August 2011). Each purple circle indicates the daily mean WVC difference (SCIAMACHY-ESA minus radiosonde) per day, where the number of measurements per radiosonde station has been used as a weight. Vertical grey lines indicate the standard deviation.
Fig. 3. Comparison between SCIAMACHY-ESA and radiosonde WVC amounts for all individual collocated cases (left) and cases where the cloud fraction is equal to zero (right) for the period of February 2010 to 10 August 2011.
Fig. 4. SCIAMACHY-ESA and radiosonde WVC differences (SCIAMACHY-ESA minus radiosonde amounts) as a function of several parameters. WVC difference as function of cloud fraction amount (top left). Green points indicate the mean WVC difference for 10 bins with an equal number of data points (5047 per bin). WVC difference as a function of cloud optical thickness (top right) (indicated in gray). Red points show the mean WVC difference for 8 bins with an equal number of data points (3873 per bin). Measurement cases where the cloud cover fraction larger or equal to 0.9 are indicated in black. WVC difference as a function of SCIAMACHY solar zenith angle (bottom left). Green points indicate the mean WVC difference for 8 bins with an equal number of data points (6308 per bin).
Fig. 5. SCIAMACHY-ESA and radiosonde WVC differences as a function of cloud top height, for all cases with a cloud cover fraction greater than or equal to 0.9. Green points indicate the mean WVC difference for 9 bins with an equal number of data points (408 per bin). Note that the two bins for the smallest cloud top heights almost overlap.
Fig. 6. Estimate of the error in the water vapour column as a function of cloud top height, as calculated with the DAK radiative transfer model for a mid-latitude summer atmospheric profile (using a total water vapour column of 2 g cm\(^{-2}\)).
Fig. 7. Global mean annual ISCCP cloud top pressure (left) and mean cloud water path amount (right) (http://isccp.giss.nasa.gov/products/browsed2.html).
Fig. 8. SCIAMACHY-ESA and radiosonde WVC differences for all individual collocated cases shown as a function of the AMF CF. Measurements where the SCIAMACHY-ESA cloud fraction are equal to zero are indicated in green, blue points indicate all other cloud fraction cases, black indicate the mean WVC difference of binned categories (AMF CF values at intervals of 0.1) and vertical black bars indicate the standard deviation from this mean.
**Figure 9.** SCIAMACHY-ESA and radiosonde mean WVC differences (SCIAMACHY-ESA minus Radiosonde WVC amounts) for the 18-month period for cloud free conditions. Each circle denotes the location of a radiosonde station and the colour indicates the mean WVC difference between the radiosonde and the collocated SCIAMACHY-ESA WVC measurements, where the number of measurements per radiosonde station has been used as a weight.
Fig. 10. SCIAMACHY-ESA and ECMWF mean WVC differences (SCIAMACHY-ESA minus ECMWF WVC amounts) for June 2005 for cloud free conditions selected according to the SPICI cloud products. Both data sets have been averaged to a ground pixel size of approximately 120 km × 30 km.
Fig. 11. Comparison between all individual collocated cases of SCIAMACHY-IMLM and radiosonde WVC amounts for the years of 2004 and 2009. The SCIAMACHY-IMLM data selection is filtered to only include land measurements taken during cloud free conditions (according to SPICI cloud product) with the additional methane column criteria applied (see Sect. 2.2.2), and excluding cases with signal level strength below 100 BU px$^{-1}$. 
Fig. 12. SCIAMACHY-IMLM WVC measurement noise error as a function of the signal strength for two years of 2004 and 2009 (note that in the figure some values in purple overlap those in black). The data used in this figure are for cloud free measurement cases with the additional methane-filter applied (see Sect. 2.2.2) and a signal level strength above 100 BUpx\(^{-1}\).
Fig. 13. The mean difference (straight line) and standard deviation (dashed line) between SCIAMACHY-IMLM and radiosonde WVC amounts as a function of the SCIAMACHY-IMLM signal strength. For the black lines the cloud-free condition is based on SPICI (700 collocations per bin), for the red lines on FRESCO+, requiring a cloud fraction < 0.02 (407 collocations per bin). SCIAMACHY-IMLM cases included in this selection are for both land and ocean measurements, excluding cases with signal strength below 100 BU px⁻¹.
Fig. 14. Comparison between collocated cases of SCIAMACHY-IMLM and radiosonde WVC amounts where FRESCO+ cloud fraction is larger or equal to 0.5 and FRESCO+ pressure larger or equal to 930 hPa, for collocated cases over land (left) and ocean (right). SCIAMACHY-IMLM measurement cases with signal strength below 100 BU px$^{-1}$ are excluded.
Fig. 15. Comparison between SCIAMACHY-IMLM and radiosonde WVC amounts for cases with FRESCO+ cloud fraction greater or equal than 0.5. The radiosonde WVC amounts are obtained by integrating the relative humidity profile from the FRESCO+ cloud top pressure, for collocated cases over land (left) and ocean (right). SCIAMACHY-IMLM measurement cases with signal strength below 100 BUpx$^{-1}$ are excluded.
Fig. 16. SCIAMACHY-ESA (top) and SCIAMACHY-IMLM (bottom) WVC amounts of 6 June 2010. The SCIAMACHY-IMLM WVC selection includes cases where the signal level strength is larger or equal to 100 BU px$^{-1}$. 
Fig. 17. Comparison between SCIAMACHY-IMLM and SCIAMACHY-ESA WVC amounts. The SCIAMACHY-ESA measurements are averaged over the SCIAMACHY-IMLM ground pixel size. Cases included in the selection are for SCIAMACHY-IMLM signal level strength is larger or equal to 100 BU px$^{-1}$ and are limited to land measurements during cloud free conditions according to the SPICI cloud parameters.