Earth System
Science
Data
Open Access

The CM SAF ATOVS data record: overview of methodology and evaluation of total column water and profiles of tropospheric humidity

N. Courcoux[1,a] and M. Schröder[1]

[1]Deutscher Wetterdienst, Satellite-Based Climate Monitoring, Offenbach, Germany
[a]now at: EUMETSAT for HE Space Operations GmbH, Darmstadt, Germany

Correspondence to: M. Schröder (marc.schroeder@dwd.de)

Abstract. Recently, the reprocessed Advanced Television Infrared Observation Satellite (TIROS)-N Operational Vertical Sounder (ATOVS) tropospheric water vapour and temperature data record was released by the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF). ATOVS observations from infrared and microwave sounders onboard the National Oceanic and Atmospheric Agency (NOAA)-15–19 satellites and EUMETSAT's Meteorological Operational (Metop-A) satellite have been consistently reprocessed to generate 13 years (1999–2011) of global water vapour and temperature daily and monthly means with a spatial resolution of $90 \, \text{km} \times 90 \, \text{km}$. The data set is referenced under the following digital object identifier (DOI): doi:10.5676/EUM_SAF_CM/WVT_ATOVS/V001. After preprocessing, a maximum likelihood solution scheme was applied to the observations to simultaneously infer temperature and water vapour profiles. In a postprocessing step, an objective interpolation method (Kriging) was applied to allow for gap filling. The product suite includes total precipitable water vapour (TPW), layer-integrated precipitable water vapour (LPW) and layer mean temperature for five tropospheric layers between the surface and 200 hPa, as well as specific humidity and temperature at six tropospheric levels between 1000 and 200 hPa. To our knowledge, this is the first time that the ATOVS record (1998–now) has been consistently reprocessed (1999–2011) to retrieve water vapour.

TPW and LPW products were compared to corresponding products from the Global Climate Observing System (GCOS) Upper-Air Network (GUAN) radiosonde observations and from the Atmospheric Infrared Sounder (AIRS) version 5 satellite data record. TPW shows a good agreement with the GUAN radiosonde data: average bias and root mean square error (RMSE) are -0.2 and $3.3 \, \text{kg m}^{-2}$, respectively. For LPW, the maximum absolute (relative) bias and RMSE values decrease (increase) strongly with height. The maximum bias and RMSE are found at the lowest layer and are -0.7 and $2.5 \, \text{kg m}^{-2}$, respectively. While the RMSE relative to AIRS is generally smaller, the TPW bias relative to AIRS is larger, with dominant contributions from precipitating areas. The consistently reprocessed ATOVS data record exhibits improved quality and stability relative to the operational CM SAF products when compared to the TPW from GUAN radiosonde data over the period 2004–2011. Finally, it became evident that the change in the number of satellites used for the retrieval combined with the use of the Kriging leads to breakpoints in the ATOVS data record; therefore, a variability analysis of the data record is not recommended for the time period from January 1999 to January 2001.

1   Introduction

Although the atmospheric CO_2 constitutes the principal "control knob" governing Earth temperature, water vapour plays a central role in the Earth's energy and water cycles by making the climate more sensitive to forcing by non-condensable greenhouse gases. In the lower troposphere, condensation of water vapour into precipitation provides latent heating which dominates the structure of tropospheric diabatic heating (Trenberth and Stepaniak, 2003a, b). Water vapour is also the most important gaseous source of infrared opacity in the atmosphere, accounting for about 60 % of the natural greenhouse effect for clear skies (Kiehl and Trenberth, 1997), and provides the largest positive feedback in model projections of climate change (Held and Soden, 2000). However, despite its great importance for climate, especially at high altitude in the tropics (Dessler and Sherwood, 2009), the behaviour and content of water vapour in the upper troposphere is not sufficiently known (Hurst et al., 2011; Kunz et al., 2013).

The Global Climate Observing System (GCOS) is a user-driven operational system intended for long-term use whose role it is to ensure availability of global observations for monitoring the climate system, detecting and attributing climate change, assessing impacts of and supporting adaptation to climate variability and change, and supporting climate research. GCOS was established in 1990 as an outcome of the second world climate conference, and it is sponsored by international and intergovernmental organisations such as the World Meteorological Organization, the Intergovernmental Oceanographic Commission, the United Nations Environment Programme, and the International Council for Science. The GCOS Second Adequacy Report (GCOS-82, 2003) established a priority list of 44 essential climate variables and called for integrated global analysis products. GCOS essential climate variables are classified into the three domains, atmospheric, oceanic, and terrestrial. Within the atmospheric domain, a distinction is made between the surface, the upper air, and the composition variables. Water vapour is one of the atmospheric surface and upper air essential climate variables because of its key role in the radiation budget, the structure of tropospheric diabatic heating, the water cycle and the atmospheric chemistry. The objective of the World Climate Research Programme's Global Energy and Water Cycle Experiment (GEWEX) is to fully understand the water cycle for predicting climate change. GEWEX has initiated a series of projects and assessments to produce long time series of parameters linked to the water cycle and to evaluate the current maturity of such products. The Global Water Vapor Project was one of GEWEX's projects dealing with water vapour, the primary goals of which were the accurate global measurement, modelling, and long-term prediction of water vapour. Furthermore, the GEWEX Data and Assessment Panel has initiated the GEWEX Water Vapor Assessment, G-VAP (http://www.gewex-vap.org). G-VAP's major objective is the characterisation of long-term satellite-based tropospheric water vapour data records, and one of its activities is the analysis of the probability density function (PDF) of water vapour.

The Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) suite onboard the TIROS-N and National Oceanic and Atmospheric Agency (NOAA)-6–14 satellites consists of three sounders – one infrared sounder, the High Resolution Infrared Radiation Sounder (HIRS), and two microwave sounders, the Microwave Sounding Unit (MSU) and the Stratospheric Sounding Unit (SSU). The MSU and SSU have since been replaced with improved instruments – Advanced Microwave Sounding Unit A and Unit B (AMSU-A and AMSU-B) – and more recently AMSU-B was replaced by the Microwave Humidity Sounder (MHS). The Advanced Television Infrared Observation Satellite (TIROS)-N Operational Vertical Sounder (ATOVS) suite, AMSU-A, AMSU-B and HIRS are onboard the NOAA-15–17 satellites. Onboard NOAA-18, NOAA-19 and Metop-A, AMSU-B has been replaced by MHS. The TOVS/ATOVS observations allow the retrieval of water vapour and temperature profiles. The TOVS/ATOVS observations started in 1978/1998 and are among the longest time series available from satellites.

Retrieval methods can be separated into statistical/semi-physical and physical schemes. The semi-physical schemes retrieve the water vapour content by applying a statistical scheme (linear regression or neural networks) based on a training data set. The physical schemes mostly use a first guess, often coming from a numerical weather forecast model or reanalysis, as the basis for the forward computation, and then vary the first-guess profile until the computed set of radiances best matches the observed radiances. Processes in the atmosphere complicate the retrieval task, e.g. the co-existence of the three thermodynamic phases of water on Earth, interaction with aerosols, and uncertainties in surface emissivities and temperatures, particularly over land. The error characteristics of the retrieval or analysis will critically depend on the a priori or training data utilised. Several retrievals for TOVS and in particular ATOVS have been developed. An important aspect in this context is that synchronised infrared and microwave observations can be used. This way the information content increases and both clear-sky and cloudy-sky conditions are sampled. An example of TOVS retrieval is described in Scott et al. (1999) and forms the basis for a data record of atmospheric profiles. Retrieval algorithms for ATOVS are described in, for example, Li et al. (2000) and Reale et al. (2008). Boukabara et al. (2011) developed the Microwave Integrated Retrieval System, which uses AMSU-A and MHS observations and is currently being updated to also include Special Sensor Microwave Imager/Sounder observations. These retrieval schemes are presently applied operationally and have not been used so far to reprocess the ATOVS record.

With the availability of hyperspectral infrared sounders which are jointly installed with microwave radiometers onboard the NASA Aqua, the EUMETSAT Metop-A/Metop-B, and the Joint Polar Satellite System's Suomi National Polar-orbiting Partnership (Suomi NPP) platforms, the retrieval capacity has been enhanced. This development started with the Atmospheric Infrared Sounder (AIRS) onboard Aqua, which has been in orbit since 2002. AIRS covers the infrared spectrum from 3.7 to 15.4 µm with a total of 2378 channels. Since 2007, EUMETSAT's Metop satellites have carried the Infrared Atmospheric Sounding Interferometer (IASI) instrument, which performs observations in the infrared spectrum (3.63–15.5 µm) with 8461 channels. Finally, the Cross-track Infrared Sounder (CrIS) onboard Suomi NPP covers the infrared spectrum (3.92–15.38 µm) with 1305 channels. Of these instruments, IASI is the only one that continuously covers the full spectral range. AIRS, IASI and CrIS retrievals are described in, for example, Susskind et al. (2011), August et al. (2014), and Gambacorta et al. (2012). Examples of evaluation results for water vapour products from ATOVS and hyperspectral instruments can be found in, for example, Bedka et al. (2010), Reale et al. (2012) and Divakarla et al. (2014).

A few long-term satellite-based water vapour profile data records have been generated and publicly released. To give an example, the NASA Water Vapor Project total precipitable water vapour (TPW) and layer-integrated precipitable water vapour (LPW) products are based on a combination of the Special Sensor Microwave Imager (SSM/I), TOVS and radiosonde data for the time period between 1988 and 1999 (Randel et al., 1996) and have contributed to the GEWEX Global Water Vapor Project. The NASA Water Vapor Project has recently been reanalysed and extended to cover the period 1988–2009 as part of NASA's Making Earth System Data Records for Use in Research Environments programme (Vonder Haar et al., 2012). An overview of available satellite and reanalysis records is provided in the G-VAP plan available at http://www.gewex-vap.org. More information on available satellite data records can also be found at http://ecv-inventory.com.

This paper introduces the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) ATOVS tropospheric humidity and temperature data record. The ATOVS observations are consistently reprocessed with a fixed processing chain. The main elements of the processing chain are the AVHRR and ATOVS Preprocessing Package (AAPP; Atkinson, 2011), the International ATOVS Processing Package (IAPP) retrieval algorithm (Li et al., 2000) and the Kriging algorithm (Schröder et al., 2013). The ATOVS data record is freely available from http://www.cmsaf.eu/wui and referenced under doi:10.5676/EUM_SAF_CM/WVT_ATOVS/V001. This paper is based on the algorithm theoretical basis document and the validation report available at http://www.cmsaf.eu/docs. After the technical specifications of the ATOVS data record, the input data are introduced, and then the preprocessing, the

retrieval and the post-processing are described. In Sect. 4, we show results from the comparison of the ATOVS data record to the GUAN radiosonde observations and the AIRS data record for the periods 1999–2011 and 2003–2011, respectively. In order to enhance readability we focus on TPW and LPW here. Finally, we provide conclusions.

2 Product description

The ATOVS data record contains tropospheric water vapour and temperature products and is defined at all longitudes and for latitudes between 80° N and 80° S. The products are available as daily and monthly means on a cylindrical equal area projection of 90 km × 90 km. The temporal coverage of the data record ranges from 1 January 1999 to 31 December 2011. The Kriging error (for daily mean products), the extra-daily standard deviation (for monthly products) and the number of valid observations per grid box are also available for each product. The data files are created following the Network Common Data Format (NetCDF) Climate and Forecast Metadata Convention version 1.5 and the NetCDF Attribute Convention for Dataset Discovery version 1.0. The products are available free of charge from the CM SAF website (www.cmsaf.eu/wui).

The following products are included in the ATOVS data record:

- Vertically integrated water vapour or total precipitable water vapour (TPW) in $kg\,m^{-2}$.

- Layered products for five layers:

    - layer vertically integrated precipitable water vapour (LPW) in $kg\,m^{-2}$,

    - layer mean temperature in K.

- Products at six pressure levels:

    - specific humidity in $g\,kg^{-1}$,

    - temperature in K.

Relative humidity for five layers is provided as additional, auxiliary data. The layer and level definitions are given in Table 1 and TPW is integrated from the surface to 100 hPa. The ATOVS data are provided on a fixed vertical grid to ease utilisation. However, the actual vertical resolution of an individual retrieval differs from pixel to pixel and time to time because the information content is a function of local surface and atmospheric conditions. The origin of the observed radiation is best described by so-called Jacobians, and in addition to atmospheric conditions these are a function of the instrument characteristics. Examples of Jacobians are given in Li et al. (2000) for AMSU-A and HIRS and in Kleespies and Watts (2006) and Buehler et al. (2004) for AMSU-B. The full ATOVS time series has been reprocessed with a fixed preprocessing, retrieval and post-processing scheme described

below. The reprocessed ATOVS data record was released in 2013. Though consistently reprocessed, the ATOVS data record may not be considered as a consistent data record, mainly because the input data require improved quality control and intercalibration.

Examples of the ATOVS data record products are shown in Figs. 1 and 2. In Fig. 1, the monthly mean TPW for September 2007 is shown together with the corresponding extra-daily standard deviation and the corresponding number of observations per grid box. Figure 2 shows LPW for the layer between 500 and 700 hPa for 27 September 2007, with the Kriging error expressed in terms of standard deviation (see Schröder et al., 2013, for a definition) and the corresponding number of observations per grid box.

Associated level 2 data are available on request. The level 2 data contain, among other information, dew point temperature on the 42 IAPP level (using the CO_2 slicing method) microwave emissivity, cloud top pressure, cloud top temperature, clear–cloudy index and effective cloud amount, total ozone, cloud fraction, rainfall, and specific humidity profiles at 42 pressure levels. However, these outputs are not part of the CM SAF ATOVS data record. The left panel of Fig. 3 shows examples of profiles of specific humidity for four different regions (Northern and Southern Hemisphere, tropics and warm pool) for September 2007. The profiles are computed as arithmetic averages over valid observations at levels smaller than or equal to the surface pressure. The specific humidity of the final product is also plotted as asterisks. The specific humidity generally decreases with height and this decrease is the strongest at 450 hPa and above. The warm pool exhibits largest specific humidity and the Northern Hemisphere is generally more humid than the Southern Hemisphere both between the surface and 200 hPa. The final product is typically more humid than the averages based on level 2 data in all regions and at all considered levels for reasons discussed in Sect. 4.2.1. The maximum difference is $0.17 \, \mathrm{g \, kg^{-1}}$ (at 850 hPa, Northern Hemisphere), which explains why the differences are hardly visible in Fig. 3.

The average differences between the ATOVS and the ERA-Interim profiles are shown in the right panel of Fig. 3. This figure illustrates the adjustment made by the retrieval to the input profiles. At near-surface layers the changes are minimal, which is likely due to the rather low information content in the observation. It is noticeable that this extends up to 650 hPa in the Southern Hemisphere. Largest reductions of up to -83% are found in the upper troposphere. While moving downward, those changes to local maxima increase by up to 11 %. These maximum values are found for the warm pool. Also shown is the difference between the final product and the input data. These differences generally exhibit very similar features to the difference between the averaged level 2 data and the input. We conclude that there are substantial changes by the retrieval in the upper troposphere and, to a lesser degree, also between 800 and 550 hPa. Whether or not these changes led to an improvement in quality can scarcely

be judged because radiosonde data are assimilated in ERA-Interim and, more generally speaking, because a true reference with sufficient spatio-temporal coverage is not available.

Finally, it should be noted here that CM SAF also provides an "operational" version of the ATOVS products with a maximum timeliness of 2 months. These data have been operational since 2009 and cover the period 2004–present. The operational processing scheme has used ECMWF Integrated Forecast System forecasts since March 2012, does not apply simultaneous nadir overpasses (SNOs) and is based on various retrieval versions. Currently, the implementation of IAPP version 4 is carried out to allow the processing of Metop-B data. The operational ATOVS products are routinely compared against GUAN observations and the results of this comparison are subject to an annual review and are published at www.cmsaf.eu/docs.

The operationally processed ATOVS data record is freely available from www.cmsaf.eu/wui.

3   Input data and retrieval

3.1   Input data

ATOVS is a sounding instrument system composed of three sounders. Two of these are microwave sounders, AMSU-A and AMSU-B, onboard NOAA-15, NOAA-16, and NOAA-17, with MHS replacing AMSU-B onboard NOAA-18, NOAA-19, and Metop-A. The third sounder is an infrared sounder, HIRS. ATOVS has been onboard NOAA and Metop polar-orbiting satellites since 13 May 1998. So far, seven platforms have carried the ATOVS instruments, namely NOAA-15–19, Metop-A, and Metop-B. AMSU-A and AMSU-B are cross-track-scanning total power radiometers with instantaneous fields of view of 3.3 and 1.1°, providing a footprint size at nadir of 48 and 16 km, respectively. The 15 AMSU-A channels primarily provide temperature sounding of the atmosphere through channels located at the 57 GHz oxygen absorption band. AMSU-A has also three channels (at 23.8, 31.4, and 89 GHz) that provide information on tropospheric water vapour, precipitation over ocean, sea ice coverage, and other surface parameters. AMSU-B has five channels that mainly measure water vapour and liquid precipitation. Three of its channels are located in the water vapour band at 183.31 ± 1, 183.31 ± 3, and 183.31 ± 7 GHz. The channels at 89 and 150 GHz are located in the atmospheric window and are sensitive to water vapour at lowest layers in the atmosphere. The MHS channels are similar to the AMSU-B channels. The third ATOVS instrument, HIRS/3 (replaced by HIRS/4 on NOAA-18, NOAA-19, and Metop-A), is an infrared 20-channel cross-track-scanning sounder with an instantaneous field of view of 1.3°, providing a nominal spatial resolution of 18.9 km (improved to 10 km for HIRS/4). HIRS infrared observations are affected by surface properties, clouds, temperature and water vapour.
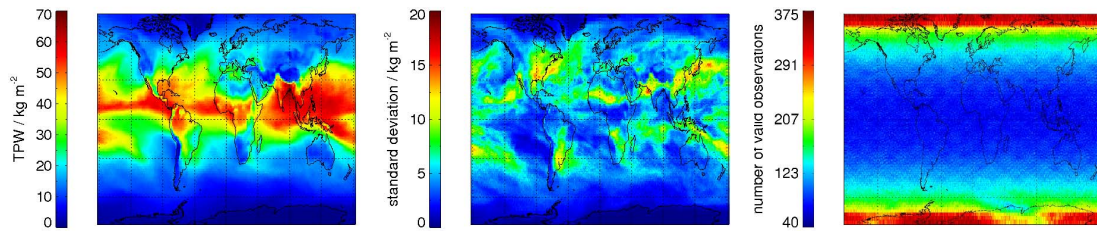
Figure 1. TPW (left panel), extra-daily standard deviation (middle panel) and number of valid observations per grid point (right panel) for September 2007.
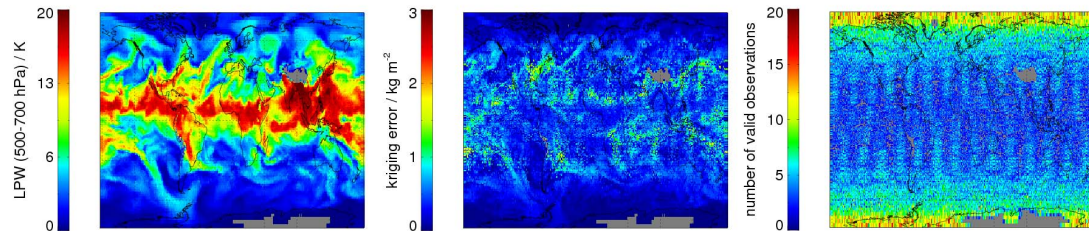

Figure 2. TPW (left panel), Kriging error (middle panel) and number of valid observations per grid point (right panel) for 20 September 2007.
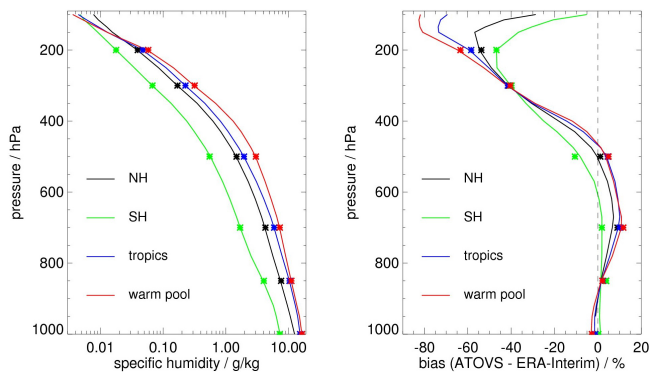

Figure 3. Average profiles of specific humidity from ATOVS (left panel) and mean difference (bias) between ATOVS and ERA-Interim (right panel) for September 2007. The regions are defined as follows: Northern Hemisphere (NH), within 20 and 50° N; Southern Hemisphere (SH), within −20 and −50° S; tropics, within −20° S and 20° N; and warm pool, within −15° S and 15° N and within 90 and 150° E. Specific humidity and bias are plotted only if the number of valid observations exceeds 75 % of the value in the upper troposphere (a minimum of 230 000 for the warm pool).

Observations from a specific satellite are used for the processing if all three ATOVS instruments are declared operational on the NOAA polar-orbiting environmental satellite status page: www.ospo.noaa.gov/Operations/POES/status.html. The number of available or operational satellites varies with time. Consequently, different combinations of satellites were used to generate the data record. Table 2 gives the details about the different satellite combinations used for the retrieval.

The retrieval of the geophysical parameters is done using IAPP software version 3.0b (see Sect. 3.3). IAPP uses the following ATOVS channels: HIRS channels 1 to 17, AMSU-A channels 1 to 15, and AMSU-B channels 17 to 20. When an instrument channel experienced a malfunction on a specific satellite, this channel was removed from the retrieval for the entire reprocessing time period for that particular satellite. Such channels are AMSU-A channels 11 and 14 on NOAA-15, AMSU-A channel 4 on NOAA-16, and AMSU-A channel 7 on Metop-A.

The IAPP relies on the use of a priori data. The following European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis fields (Dee et al., 2011) are used as a priori information: temperature profile, relative humidity profile, 2 m dew point, 2 m temperature, skin temperature, surface pressure, geopotential height, sea ice cover, land–sea mask, and total column water vapour.

3.2 Input data preprocessing

The input data preprocessing is carried out in two steps. First, AAPP is used to convert the geo-referenced and calibrated brightness temperatures (level 1c, taken from ECMWF's Meteorological Archival and Retrieval System) into mapped data (level 1d). During this process the scan lines are also sorted according to time. Furthermore, the AAPP Binary Universal Form for the Representation of Meteorological Data decoding tool is used to read the l1c data. The AAPP software is developed and maintained by the EUMETSAT Satellite Application Facility for Numerical Weather Prediction. An overview of AAPP is given in Atkinson (2011), a scientific description is available from Labrot et al. (2011), and the software description can be found in Labrot et

Table 1. Layer and level definitions for the ATOVS data record.

Layer	1	2	3	4	5	–
Pressure [hPa]	300–200	500–300	700–500	850–700	Surface–850	–
Level	1	2	3	4	5	6
Pressure [hPa]	200	300	500	700	850	1000

Table 2. Satellite combinations used to generate the ATOVS humidity and temperature data record together with the corresponding time period.

Time period (yyyy/mm/dd)	Satellite used
1999-01-01–2000-10-31	NOAA-15
2000-11-01–2001-01-31	NOAA-16
2001-02-01–2002-10-31	NOAA-15, NOAA-16
2002-11-01–2003-09-30	NOAA-15, NOAA-16, NOAA-17
2003-10-01–2005-08-31	NOAA-15, NOAA-16
2005-09-01–2007-05-31	NOAA-15, NOAA-16, NOAA-18
2007-06-01–2009-01-31	NOAA-15, NOAA-16, NOAA-18, Metop-A
2009-02-01–2009-04-30	NOAA-15, NOAA-16, Metop-A
2009-05-01–2009-06-30	NOAA-16, Metop-A
2009-07-01–2011-12-31	NOAA-16, Metop-A, NOAA-19

al. (2012). The default AAPP version was used. The HIRS pixel definition defines the "grid" for AAPP preprocessing.

Secondly, SNO coefficients are applied to the data of the four AMSU-B channels used for the retrieval (channels 17 to 20) to intercalibrate observations from the different satellites. The SNO coefficients used to process the ATOVS data record are described in John et al. (2012) and were provided (V. John, personal communication, 2010) as monthly mean brightness temperature differences for the satellites NOAA-15 to NOAA-18 and Metop-A, covering the period January 2001–December 2010. Since NOAA-16 exhibits temporal overlap with all other satellites that have ATOVS instruments onboard, it has been used as a reference satellite for the SNO intercalibration. John et al. (2012) emphasise that the quality of the intercalibration using classical SNO approaches is hampered due to the overrepresentation of cold scenes. The biases between the satellites are dependent on the scene radiance, which is itself dependent on the latitude at which the observation is made. Improvements to classical SNO approaches were suggested by John et al. (2012) for AMSU-B and developed by Shi and Bates (2011) for HIRS. Unfortunately, at the time of the data record processing, no intercalibration coefficients based on the conclusions of John et al. (2012) were available. In general, intercalibration coefficients are also available for AMSU-A (see Zou and Wang, 2011, for details) and HIRS (see Shi and Bates, 2011, for details). However, they are applicable to limb-corrected observations and thus not useable for the processing of the ATOVS data record as IAPP requires non-limb-corrected radiances as input. Consequently, intercalibration coefficients have not

been applied to the HIRS and AMSU-A data for the processing of the CM SAF ATOVS record.

3.3 Retrieval

The retrieval software used to generate the ATOVS data record is IAPP version 3b developed by the University of Wisconsin in Madison, WI, USA (Li et al., 2000). The default version of IAPP was used, as no parameters can be tuned in the IAPP configuration file, which mostly contains path definitions for the different data needed for the retrieval. The IAPP retrieves, among other atmospheric parameters, temperature and moisture profiles in both clear and cloudy atmospheres at 42 pressure levels. The IAPP algorithm can be decomposed into the following steps: the HIRS cloud detection and removal procedures, the bias adjustment relative to collocated radiosonde observations, and the actual retrieval. The bias adjustment scheme is applicable to NOAA-15 data only. It has not been applied here because it has been anticipated that its application will lead to a breakpoint in the time series of the final products. The goal of a bias correction is to account for calibration uncertainties of the satellite data, radiative transfer uncertainties and uncertainties of the input to the radiative transfer. The deactivated bias correction can impact the number of convergent retrievals and the systematic and random uncertainties of the retrieved parameters. The retrieval involves two steps. In the first, the initial temperature, water vapour, ozone profiles, and the surface skin temperature are obtained by statistical regression between the ATOVS measurements and the ERA-Interim reanalysis. The second part of the retrieval is the computation of an iter-

ative physical solution of the radiative transfer equation using the first-guess results and the ERA-Interim reanalysis as background information. The physical iterative retrieval algorithm, the cloud detection procedure and the bias adjustment method are described in detail in Li et al. (2000) and are reiterated in Courcoux and Schröder (2013). Here, we note that the HIRS cloud detection algorithm is applied to 3×3 adjacent HIRS pixels. When one or more pixels are cloud-free, the retrieval process is applied. If this is not the case, the cloud removal process can be applied; however, it is not implemented in IAPP version 3b utilised here. A land-only and ocean-only scattering index threshold is applied to AMSU-A observations in order to flag pixels affected by strong scattering events, which typically occur in the presence of strong precipitation or in the presence of snow cover and sea ice. The microwave surface emissivity is part of the solution, while the infrared surface emissivity is set to 0.99 during the retrieval process (Li et al., 2000).

3.4 Post-processing

The retrieval outputs are first quality-controlled according to the following criteria:

- TPW between 0 and $90 \, \mathrm{kg \, m^{-2}}$,

- temperature between 180 and 340 K,

- specific humidity between 0 and $55 \, \mathrm{g \, kg^{-1}}$,

- surface emissivity between 0 and 1,

- surface pressure between 0 and 1050 hPa (on basis of input data).

If profile or surface values are outside these ranges or if the profile exhibits super-adiabaticity, the full profile is set to undefined. After quality control, the 42 level profiles are integrated and averaged to obtain the final products described in Sect. 2.

Finally, an objective interpolation technique commonly called Kriging is applied to the quality-controlled and integrated products. The advantage of applying Kriging is that it fills data gaps and that uncertainty estimates at grid level are computed. The principle of Kriging is that an estimate or prediction for an unobserved location is computed by using the observations from locations in its vicinity. The optimal estimate at each grid point is found by a weighted average of the information from the surrounding points. The challenge is to determine these optimal weights. The weights depend on the distance-dependent spatial correlation function and the error of the observation used. The Kriging algorithm used for the ATOVS data record is described in detail in Schröder et al. (2013). The only parameter tunable by the user in the Kriging algorithm is the grid resolution – here up to $90 \, \mathrm{km} \times 90 \, \mathrm{km}$. The extra-daily standard deviation for the monthly means, the Kriging error for the daily means,

and the number of valid observations per grid box, which are outputs of the Kriging algorithm, are part of the ATOVS data record.

4 Evaluation

The ATOVS tropospheric humidity and temperature data record is compared to GUAN radiosonde observations in order to guarantee consistent and comparable evaluation results between the operational and the reprocessed ATOVS data records. To further allow a global comparison we also use the AIRS data record. AIRS observations have a large temporal overlap with the ATOVS data record. Many other ground-based, in situ and satellite observations are available for comparison. An extensive list of such data records is given in the appendix of G-VAP plan, available at www.gewex-vap.org.

The goal of the comparison of the ATOVS data record with GUAN radiosonde and AIRS data record is to identify and understand potential issues in the ATOVS data record and to provide an overall characterisation of the ATOVS data record in a relative sense. An accuracy assessment is not carried out. Furthermore, the impact of background information and uncertainty on the observed quality is not analysed here, and we refer the reader to, for example, Eyre and Hilton (2013) for further reading.

In Sect. 4.1 the GUAN and AIRS data records are described. The comparison considers TPW and LPW and the results are presented in three subsections of Sect. 4.2. In the first, the TPW time series from ATOVS, GUAN and AIRS data records are presented and discussed. In the second and third, the comparison results between ATOVS and GUAN data records and between ATOVS and AIRS data records are discussed.

4.1 Data for evaluation

4.1.1 GUAN

The GUAN radiosonde network has been established by GCOS in order to make current and historical upper air data available for climate change detection and climate monitoring. GUAN provides global radiosonde observations, from homogeneously distributed upper air stations, that have a specific record length in addition to meeting the continuity requirement and data quality requirements as defined by GCOS (Daan, 2002). At present there are 171 GUAN stations worldwide. A station map and a station list can be found at http://www.wmo.int/pages/prog/gcos/index.php?name=ObservingSystemsandData. The GUAN data are distributed by the Global Telecommunication System and archived at the Deutscher Wetterdienst. The processing of GUAN data at the Deutscher Wetterdienst was consistently done, with one exception: in October 2008 the archiving system at the Deutscher Wetterdienst was changed and the GUAN processing software was adapted. However, the re-

sults of the comparison between ATOVS and GUAN data records do not exhibit any distinct feature in October 2008.

The quality of radiosonde observations is affected by a series of issues such as temporally and spatially varying radiosonde types and national practice (e.g. Soden and Lanzante, 1996; Christy and Norris, 2009; Moradi et al., 2010), as well as issues and differences in calibration procedures (e.g. Miloshevich et al., 2006; Vömel et al., 2006). Among the strongest impacts is the dry bias caused by solar radiation (Vömel et al., 2006), which leads to significant underestimations of humidity in the upper troposphere if not corrected. A series of correction algorithms have been developed by, for example, Miloshevich et al. (2004), Leiterer et al. (2005) and Miloshevich et al. (2009), which mainly focus on RS80 and RS92 radiosonde observations. Such corrections have not been applied to the utilised GUAN observations.

Examples of reprocessed radiosonde archives which include temperature and water vapour are the integrated global radiosonde archive (Durre et al., 2006) and its homogenised version (Dai et al., 2011). Dai et al. (2011) describe a few known discontinuities in humidity observations from radiosondes. These are as follows: the dew point depression was set to 30 °C under dry conditions at several stations, and temperature observations under cold conditions for "early radiosonde hygrometers" were unreliable and were reported as missing (Dai et al., 2011).

4.1.2   AIRS

AIRS is an infrared cross-track-scanning instrument onboard the NASA Aqua satellite which also carries an AMSU-A radiometer. The NASA Aqua satellite has been in orbit since 2002. The level 2 AIRS data record which is used for comparison is the AIRX2RET product provided by the NASA Goddard Earth Science Data and Information Service Center (http://daac.gsfc.nasa.gov/); this product is based on AIRS and AMSU-A observations. The processing version is V5.0 for the data from 2002 to 30 September 2007 and V5.2 for data from 1 October 2007 onwards. AIRS L2 products come in swath-based 6 min length files, with 240 files covering 1 day. The products have a spatial resolution of $50 \, \text{km} \times 50 \, \text{km}$ and the profiles are defined on 14 layers at 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70 and 50 hPa. The format is HDF4. The AIRS Standard Products consist of, among other things, cloud properties and profiles of temperature and water vapour. The products are the results of employing the combined AIRS-IR/AMSU-A microwave retrieval, which is described in detail in Susskind et al. (2003, 2006, 2011). The retrieval process also involves a cloud-clearing process which assumes that the radiative properties in each field of view are a function of cloud fraction only. A retrieval solution is rejected when the cloud fraction is larger than 90 %. A scattering (rain) index is not explicitly applied. Infrared and microwave surface emissivities are part of the solution. For the comparison to the ATOVS data record, the data field Qual_H2O was evaluated and "best" and "good" quality data were used in the comparison.

An evaluation of the AIRS version 5 TPW products' accuracy is given in Bedka et al. (2010), who compared the satellite products to ground-based observations at selected Atmospheric Radiation Measurement (ARM) sites. Using ground-based microwave radiometer observations at Barrow, Southern Great Plains–Lamont (SGP) and Nauru, the authors found an average relative error which is typically smaller than 5 % for all sites, except at SGP, where AIRS products are too moist when TPW is less than $10 \, \text{kg} \, \text{m}^{-2}$. At SGP, night-time observations exhibit a dry bias of approximately 10 % when TPW is greater than $10 \, \text{kg} \, \text{m}^{-2}$ (Bedka et al., 2010).

Recently, the AIRS version 6 products have been released. Improvements over version 5 are described at http://airs.jpl. nasa.gov/data/v6/; the first comparison results of version 6 and version 5 products to ECMWF can also be found on the webpage. The AIRS version 6 products may still be improved, and product validation and quality assurance are ongoing. Thus, its product validation state is "provisional".

4.2   Evaluation results

4.2.1   Time series of the different data records

Three data records are used for the evaluation: ATOVS and GUAN data records for the period 1999–2011 and AIRS data record for the period 2003–2011. The data have not been collocated and GUAN data have only been used when at least two observations per day are available. ATOVS and AIRS data records exhibit similar annual cycles. However, a systematic difference between both data records is evident. This is discussed in Sect. 4.2.3. The annual cycle of the GUAN time series has larger amplitudes than the annual cycles of the satellite time series (not shown). The GUAN stations are located on islands and over land, with the majority of stations in the continental Northern Hemisphere. Schröder and Lockhoff (2013) show that the strength of the annual cycle is a function of region: strongest annual cycles are associated with the monsoon regions and the propagation of the ITCZ, largest regions of minimum strength are found over the oceans of the Southern Hemisphere, and land areas typically exhibit strong annual cycles. The former explains the presence of an annual cycle in the satellite data due to the imbalance in strength between the Northern and the Southern Hemisphere and the latter in combination with the asymmetric sampling between the Northern and Southern Hemisphere explains the annual cycle in the GUAN data. The annual cycle in TPW from GUAN has slightly larger amplitudes in 1999 and 2000 than from 2001 onwards (not shown). The larger amplitudes in the first 2 years are caused by stronger minima during boreal winters. When looking at the time series of deseasonalised anomalies (not shown), a breakpoint is found between June and July 2001. The strength of the breakpoint is computed as the difference between the aver-
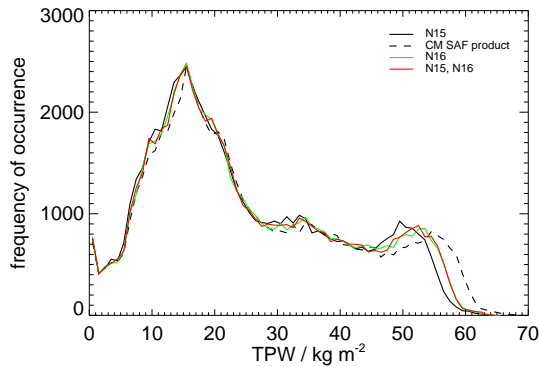
Figure 4. Histogram of the TPW values for June 2002. The dashed line represents the histogram of the CM SAF ATOVS TPW product using the data from the NOAA-15 and NOAA-16 satellites and the Kriging method for averaging. The red line represents the histogram of the data from the NOAA-15 and NOAA-16 satellites being averaged using the arithmetic averaging method. The solid black and green lines represent the data from the NOAA-15 and NOAA-16 satellites, respectively (averaged also using the arithmetic averaging method).

age TPW using a period of 24 months prior to and after the break. Through use of this strength and the averaged standard deviation over the two periods as input to a two-sided t test, it is found that the strength of the change is associated with a coverage probability of 92 %. Thus, the breakpoint is not considered to be significant when applying a standard significance level of 0.05. In the following the strength and the coverage probability are computed and interpreted the same way.

The ATOVS TPW data record exhibits a breakpoint between January 2001 and February 2001 (not shown). The difference in TPW between the years 1999–2000 and 2002–2003 is 2.8 kg m^{-2} with a coverage probability of 99 %. Moreover, the annual cycle of TPW exhibits stronger minima during boreal winters during the first 2 years. This breakpoint is analysed in more detail in the following.

First, the breakpoint does not temporally coincide with the start of the use of SNOs in January 2001. Moreover, no breakpoint is visible between December 2010 and January 2011, when the use of SNOs ends.

Second, we assess the impact of Kriging on the homogeneity of the time series. We compared the PDF based on the CM SAF ATOVS data record products and ATOVS products, which have been arithmetically averaged on basis of daily values. Figure 4 shows the PDFs of TPW values for June 2002 separately for the CM SAF ATOVS data record products and for the arithmetically averaged monthly means. Obviously, the distribution of the CM SAF ATOVS data record products exhibits an increased number of TPW values at the high end of the distribution. This is reflected in the monthly mean TPW of 25.3 kg m^{-2} for the classically averaged data and of 26.2 kg m^{-2} for the CM SAF data record

products, which gives an average difference of 0.9 kg m^{-2}. Apart from sampling, gaps are caused by strong scattering events, e.g. in the presence of strong precipitation. During gap filling, Kriging uses valid observations in the vicinity of the gaps. The gaps neighbouring areas are typically humid, and thus Kriging fills these gaps with generally large values (see also Schröder et al., 2013). This explains the increased frequency of occurrence at the high end of the TPW distribution. The PDF does not change significantly when more than two satellites are used (not shown). The PDFs of the classically averaged monthly means for NOAA-15 and NOAA-16 alone are also shown in Fig. 4. The difference between the arithmetically averaged monthly means for NOAA-15 and NOAA-16 is 24.7 and 25.2 kg m^{-2}, respectively, thus leading to a difference of 0.5 kg m^{-2}.

Finally, a specific feature of Kriging is discussed. Kriging requires two independent measurements such as those from different satellites or from the morning and afternoon overpasses of a single satellite. For the period January 1999 to February 2001, only NOAA-15 was available. Then, it may happen that the same location is not observed twice a day, e.g. due to the occurrence of strong precipitation events. When this happens, Kriging is not applied and the daily average is flagged as undefined. For the June 2000 case study, the number of valid observations is 12 % smaller in the Kriging product than in the arithmetically averaged product and the number of undefined values is 9 % larger in the Kriging product than in the arithmetically averaged product. Indeed, it is visible that the positions of minima in the number of valid observations and of undefined values coincide with the position of the Intertropical Convergence Zone (ITCZ) (not shown).

4.2.2 Comparison to GUAN data

The methodology for the comparison of the ATOVS data record against the GUAN radiosonde data record is as follows. First, the GUAN data record is integrated to match the vertical layer and level definitions of the ATOVS data record water vapour products. For each day, only stations with at least two radiosonde launches per day are used and averaged to daily values. The ATOVS data record is spatially collocated to the position of the GUAN stations using a nearest-neighbour algorithm. The collocated daily averages form the basis for the comparison. We analyse the monthly bias and the bias-corrected root mean square error (RMSE) between ATOVS and GUAN data records. The number of valid collocations per month is greater than or equal to 450. The results shown in this section show bias and RMS based on all valid daily averages. Note that potential dependencies on climate regimes, TPW, and other regional dependencies are not resolved here. We expect occasionally larger bias and larger RMS on regional scale.

First of all, the difference in quality between the reprocessed and operational ATOVS products is discussed. Fig-
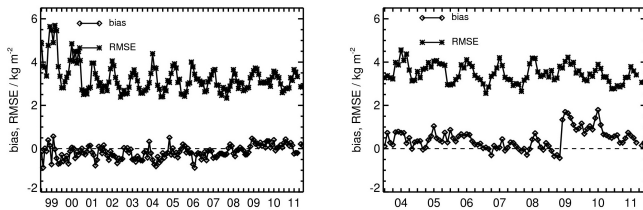
Figure 5. TPW bias and bias-corrected RMSE between the ATOVS and the GUAN data records: reprocessed data set from January 1999 to December 2011 (left panel) and operational data set from January 2004 to December 2011 (right panel). Note the difference in temporal coverage.

ure 5 presents the comparison results between TPW from the reprocessed and operational ATOVS data records and the GUAN data record. Figure 5 clearly shows that the TPW product from the reprocessed ATOVS data record exhibits a better quality and stability than the TPW from the operational ATOVS product. The bias of the operational ATOVS product compared to the GUAN data record shows a significant breakpoint between April and May 2009. At this time the following changes had been implemented in the operational ATOVS processing chain: migration of the processing chain, update of AAPP and IAPP, removal of NOAA-15 and NOAA-18 observations from the retrieval, and implementation of Metop-A and NOAA-19 observations in the retrieval. The obvious improvement for the reprocessed data record is that the breakpoint in the bias time series is largely reduced, and this also leads to an improved averaged bias. See Sect. 4.2.3 for further discussion. Moreover, the RMSE is slightly smaller for the reprocessed data record than for the operational product.

We now focus on the comparison between the reprocessed ATOVS data record and the GUAN data record that is shown in the left panel of Fig. 5. The TPW bias is typically smaller than $1 \, \mathrm{kg \, m^{-2}}$ and the average bias is $-0.16 \, \mathrm{kg \, m^{-2}}$; furthermore, the bias is stable. Interestingly, no breakpoint is observed in the bias time series between 2000 and 2001. The breakpoints in the averaged, non-collocated TPW time series of GUAN and ATOVS have the same direction and a different strength and occur with a temporal difference of 5 months. This might translate into a breakpoint in the bias time series which is smaller than the ATOVS breakpoint itself and which is overlain by an anomaly between February and June 2001. This is not evident in the bias time series due to the collocation process: gap filling is mainly applied in the presence of strong precipitation. Associated areas are typically found in the ITCZ and storm track regions with poor coverage of GUAN stations. Due to the collocation procedure, data from these areas have a reduced impact on the breakpoint. In fact, the ATOVS time series of collocated TPW values exhibits a breakpoint which is reduced by $0.7 \, \mathrm{kg \, m^{-2}}$. The GUAN data record and the ATOVS data record are sampled in a similar way; consequently, the collocated GUAN data also exhibit
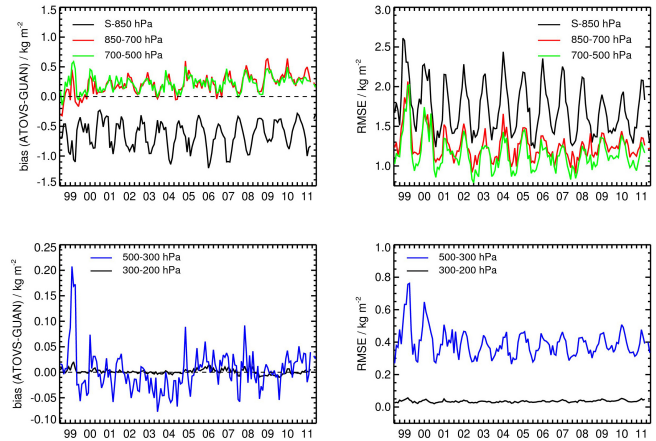

Figure 6. LPW bias (left panel) and bias-corrected RMSE (right panel) between the ATOVS and the GUAN data records for the time period January 1999 to December 2011. The upper panels show the time series for the three lowermost layers and the lower panels show the time series for the two uppermost layers.

a breakpoint between January and February 2001, and the breakpoint is not evident in the bias time series.

The averaged RMSE is $3.25 \, \mathrm{kg \, m^{-2}}$ and the RMSE is stable from 2001 onwards, with values around $3 \, \mathrm{kg \, m^{-2}}$. The RMSE exhibits maximum values in the first 2 years of the time series. This is expected since, for the years 1999 and 2000, only one satellite is available for the processing, while for the rest of the processing at least two satellites are used. This behaviour was also observed in a comparison between SSM/I-based and ERA-Interim TPW products: the RMSE decreased with the transition from a single-satellite product to a multiple-satellite product (Schröder et al., 2013). Finally, in contrast to the bias, the RMSE exhibits an annual cycle with maxima during boreal summers. When analysing the standard deviation of TPW from GUAN data record (not shown) a pronounced annual cycle is visible with a sharp increase in standard deviation between May and July (maximum value: $4.3 \, \mathrm{kg \, m^{-2}}$). The maximum in July is followed by a slow decrease until February (minimum value: $2.3 \, \mathrm{kg \, m^{-2}}$). Besides a potential dependency of the uncertainty of the radiosonde observations on TPW, we argue that the dominating factor for the annual cycle in RMSE is that, with increasing TPW values during boreal summers, the natural variability in water vapour also increases and that the increase in natural variability enhances the representativity uncertainty between the point and areal observations.

Figure 6 presents the comparison results between the ATOVS and the GUAN LPW products, again in terms of bias and RMSE. The LPW bias for layer 5 (surface–850 hPa) is around $-0.7 \, \mathrm{kg \, m^{-2}}$, while the LPW biases for layers 3 and 4 (850–700 and 700–500 hPa) are between 0 and $0.6 \, \mathrm{kg \, m^{-2}}$. As for the TPW, a slight increase in bias is present in early 2009, and an explanation for this increase is given in

Sect. 4.2.3. The LPW biases for layers 1 and 2 are relatively small, with values around 0.002 and 0.003 kg m^{-2}, respectively. The LPW bias for layer 2 exhibits an unexplained anomaly of approximately 0.2 kg m^{-2} in 1999. Maximum RMSE values are slightly larger than 2.5 kg m^{-2} and are found for layer 5. LPW RMSE values for layers 2 to 4 exhibit a decrease after the first 2 years, as was found for TPW. Furthermore, the RMSE typically exhibits an annual cycle similar to the TPW RMSE.

In view of the results shown in Fig. 3 we briefly want to characterise the relative bias of the ATOVS specific humidity product (not shown). The relative bias increases with height and ranges from 4 % at 1000 hPa to approximately 90 % at 200 hPa, with this latter value being 3 or more times larger than the values of the other levels and with ATOVS being more humid than the radiosondes. This may again partly be explained by the dry bias in radiosondes. However, the relative bias is of similar order to the maximum values given in Fig. 3. This may point to a wet bias in the ATOVS product in the upper troposphere. However, a verification is hard to accomplish due to the lack of fully independent and high-quality reference data.

When relative values are considered (not shown), the bias is the smallest (largest) for LPW5 (LPW1) with average values of -6 and 36 %, respectively. The moist bias in the upper troposphere can be expected due to the observed dry bias in uncorrected radiosonde observations (e.g. Soden and Lanzante, 1996). The relative RMSE systematically increases with height. The lowest layer has an average relative RMSE of 30 %, whereas for the highest layer this is 107 %. Our results exhibit similar magnitudes to the results presented in Reale et al. (2012), who compared humidity profiles from various satellite products, among them the NOAA ATOVS product (Reale et al., 2008), with the Global Telecommunication System radiosonde data and found typical values of 25–50 % below 400 hPa and of 100 % or more in the upper troposphere.

The GUAN radiosonde data and ATOVS are assimilated in the ERA-Interim reanalysis. Consequently, the bias and RMS of the comparison between the ATOVS data record and the GUAN radiosonde data might be underestimated due to this dependency. Although it is outside of the focus of this work, we briefly note again that various uncertainties contribute to the observed differences. Here, we compare point measurements with areal observations. Thus, the representativity uncertainty impacts the observed differences. To our knowledge the representativeness uncertainty is not known at each GUAN station. However, for assimilation purposes, high-resolution models have been used to assess such uncertainties. An analysis example is given in Waller et al. (2014) for specific humidity – they found a strong dependence of the representativity uncertainty on height and weather state. Furthermore, the comparison of the ATOVS and the GUAN data is based on daily averages and the differences in sampling between the radiosonde and the satellite observations contribute
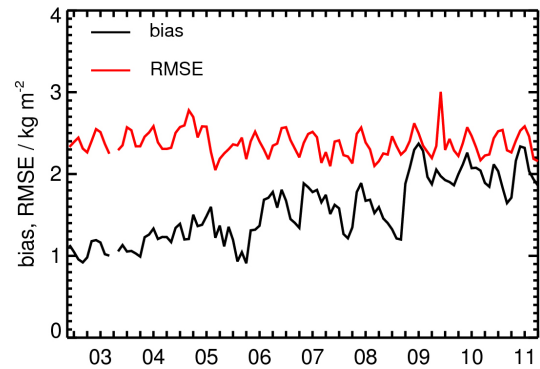

Figure 7. TPW bias and bias-corrected RMSE between the ATOVS and the AIRS V5 data records for the time period January 2003 to December 2011.

to the observed differences. To give an example of the diurnal sampling uncertainty, we use the work of Dai et al. (2002). Using high-temporal-resolution Global Positioning System data from stations over North America, they found that the uncertainty in seasonally averaged TPW is within ± 3 % or ± 0.5 kg m^{-2} when the sampling is changed from 30 min to twice daily at 00:00 and 12:00 UTC. Finally, we refer the reader to the work of Pougatchev et al. (2009) and Sun et al. (2010), who compared temperature and relative humidity profiles from radiosonde data to IASI and to the Constellation Observing System for Meteorology, Ionosphere, and Climate observations in order to analyse the uncertainty arising from temporal and spatial mismatches.

4.2.3 Comparison to AIRS data

In order for it to be possible to compare the ATOVS data record with the AIRS data record, the AIRS profiles are vertically integrated according to the ATOVS layer definitions (see Sect. 2). Then, the swath-based products are gridded onto the ATOVS spatial grid, and finally all data are averaged to obtain monthly means, which form the basis for the comparison. The number of valid collocations per month is typically larger than 60 000.

Figure 7 presents the comparison results of AIRS and ATOVS TPW products. It can be seen that the TPW bias changes from approximately 1 kg m^{-2} in 2003 to approximately 2 kg m^{-2} in 2011. A breakpoint is present between April and May 2009 which temporally coincides with the removal of NOAA-15 data from ATOVS processing. The strength of the breakpoint is 0.5 kg m^{-2} and exhibits a coverage probability of 97 %. The RMSE is relatively stable, with values around 2.4 kg m^{-2}.

The LPW bias is shown in Fig. 8 and exhibits similar features as the bias for TPW, except for the LPW bias for layer 5, which exhibits an annual cycle. The breakpoint observed in the comparison of TPW in early 2009 is also evident for LPW for layers 3 to 5. Relative to the TPW bias and the
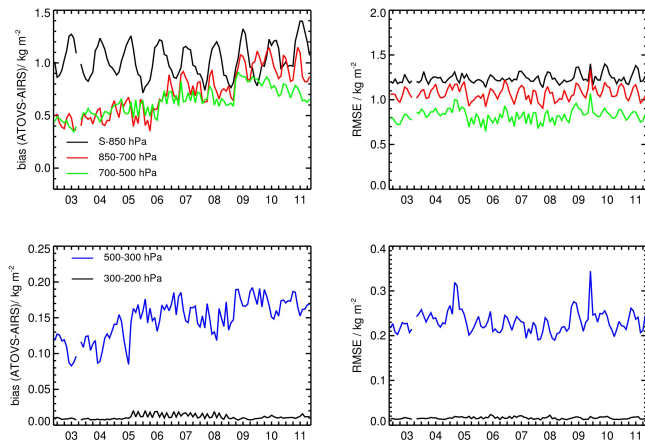
Figure 9. Mean TPW bias between ATOVS and AIRS V5 data records for the time period January 2003 to December 2011.

Figure 8. Similar to Fig. 6 but for the bias and bias-corrected RMSE between the ATOVS and the AIRS V5 data records for the time period January 2003 to December 2011.

LPW bias for layers 3 to 5, the RMSE is stable over time. The LPW bias for layer 1 and the LPW RMSE for layer 3 exhibit a distinct feature between late 2005 and early 2009. This coincides with the use of NOAA-18 observations in the retrieval while MHS onboard this particular satellite experienced a series of technical issues (see http://www.ospo.noaa.gov/Operations/POES/NOAA18/mhs.html).

The RMSE between the ATOVS and the AIRS data records does not exhibit a pronounced annual cycle and is typically smaller than the RMSE between the ATOVS and the GUAN data records likely because the number of valid collocations is larger and equally distributed over the Northern and Southern Hemisphere and because the comparison of point measurements with areal observations likely exhibits larger representativity uncertainties than the comparison of two areal observations. However, the biases for TPW and LPW are larger between the ATOVS and the AIRS data records than between the ATOVS and the GUAN data records. The relatively large bias between the ATOVS and the AIRS data records is discussed and analysed in more detail. Schneider et al. (2012) compared the TPW of a SSM/I+Medium Resolution Imaging Spectrometer (MERIS) product to TPW products from GUAN, AIRS (V5) and ATOVS (operational CM SAF ATOVS data) for the period 2004–2008. For the comparison to AIRS, the AIRS cloud mask has been applied because TPW from MERIS is only available under clear-sky conditions. They found biases (RMSEs) of 0.5 and $-1.1\,\mathrm{kg\,m^{-2}}$ (2.3 and $3.3\,\mathrm{kg\,m^{-2}}$) relative to AIRS and ATOVS, respectively. Due to the clear-sky bias (e.g. Sohn and Bennartz, 2008; Mieruch et al., 2010), these values cannot be directly compared to the results of this work. Nevertheless, this study shows that the ATOVS product is more humid than the AIRS product, and the comparison of the ATOVS data to the AIRS data exhibits similar RMSE values to our analysis. Schneider et al. (2012) also
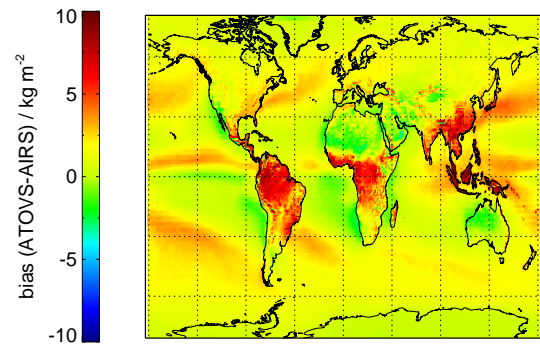
observed a greater bias in their comparison to the AIRS data record relative to their comparison to GUAN data record.

Bedka et al. (2010) compared the AIRS V5 data record to ARM observations at Nauru, Barrow and SGP. The average RMSE values are between 2.0 and $3.4\,\mathrm{kg\,m^{-2}}$ and envelope the average RMSE of $2.4\,\mathrm{kg\,m^{-2}}$ observed here. The bias between AIRS and ARM data is on average smaller than $0.1\,\mathrm{kg\,m^{-2}}$. However, their comparison results to all sites exhibit a day–night contrast which is most pronounced for SGP, with day–night differences of $1.6\,\mathrm{kg\,m^{-2}}$. They conclude that the relative difference at SGP for values larger than $10\,\mathrm{kg\,m^{-2}}$ is 10 % during boreal summers and decreases during boreal winters. They argue that surface emissivity, land use and cover, and unique boundary layer conditions may contribute to this difference.

Finally, Fig. 9 shows a map of the TPW bias between ATOVS and the AIRS data records. Obviously the bias is dominated by regions of strong precipitation and frequent cloud occurrences such as in the ITCZ and storm track regions. Relatively large values are observed at mountainous areas such as the Alps, but the maximum differences are observed over tropical land surfaces with relative differences of about 15 %. Of course, differences in retrieval setup and associated uncertainties contribute to this bias. Of particular relevance in view of the spatial distribution of the bias are differences in cloud detection, in cloud clearing (not applied for the ATOVS data record) and in the handling of scattering events (screened in the ATOVS data record). In the ATOVS data record, AMSU-B observations are used which also allow a retrieval under cloudy conditions, while in the AIRS data record, cloud clearing needs to be applied to the AIRS data in order to retrieve TPW. In general clear-sky observations exhibit a systematic underestimation of TPW relative to almost all-sky observations (e.g. Sohn and Bennartz, 2008). Thus, the different instrumentation might contribute to the observed differences.

As outlined earlier, the gap-filling process of the Kriging contributes to the observed difference between the ATOVS and the AIRS TPW products with the TPW from ATOVS be-

ing larger in precipitating areas than the TPW from AIRS. Schröder et al. (2013) compared the CM SAF SSM/I TPW product to the SSM/I TPW product from the University of Hamburg and from the Max Planck Institute for Meteorology (Andersson et al., 2010). The only difference in the generation of both products is again that the CM SAF product is based on post-processing using Kriging. The spatial distribution of their results is very similar to spatial distribution in Fig. 9. Thus, Kriging is a significant contributor to the observed bias between the ATOVS and the AIRS data records.

As precipitation over tropical land surfaces exhibits a pronounced diurnal cycle with maxima in the late afternoon and evening (e.g. Yang and Slingo, 2001) the differences between TPW from Metop-A (Equator-crossing time of ascending node: $\sim$ 21:30 local time), NOAA-16 ($\sim$ 19:00 local time) and NOAA-19 ($\sim$ 13:30 local time) have been analysed for the year 2011 (not shown). While the differences between the TPW from NOAA-19 and from Metop-A are relatively small, the differences between the TPW from NOAA-16 and from the other two satellites exhibit pronounced minima over tropical land surfaces. We found minimum differences of approximately $-3\,\mathrm{kg\,m^{-2}}$, or $-2\,\%$, with smaller TPW values from NOAA-16 than from NOAA-19 and from Metop. These minima over tropical land surfaces nearly vanish when NOAA-16 and Metop-A differences are computed for the year 2008. Then, the NOAA-16 Equator-crossing time is approximately 16:00 local time. It seems that the diurnal sampling in combination with the diurnal cycle of deep convection over tropical land surfaces has an impact on the ATOVS product. However, channel 4 observations from NOAA-16 have not been used as input to the retrieval. Information from channel 4 is valuable to separate information on near-surface properties from the temperature and water vapour signal of the lower troposphere. When the differences in the TPW products from the CM SAF ATOVS data record and from the AIRS data record are compared for the months of July and August for the period between 2008 and 2011 (not shown), the general coincidence of large biases and precipitating areas is still present. However, land surfaces in the northern extratropics exhibit an increase in bias as well, which might also be associated with convective precipitation. Because convective precipitation has a short-term impact on surface emissivity, differences in handling surface emissivities contribute to the overall difference as well.

Because the TPW bias between the ATOVS and AIRS data records exhibits a breakpoint in early 2009, we had a closer look at the TPW time series from the ATOVS data record and at the TPW bias time series between the ATOVS and the GUAN data records. Between April and May 2009, the ATOVS TPW anomaly time series exhibits a breakpoint with a strength of $0.75\,\mathrm{kg\,m^{-2}}$ with coverage probability of 98 % and the bias between TPW from ATOVS and from GUAN, a breakpoint of strength $0.37\,\mathrm{kg\,m^{-2}}$ with coverage probability of 97 %. Obviously the removal of NOAA-15 observations from the retrieval in May 2009 introduces a breakpoint

into the ATOVS time series. After the removal of NOAA-15 observations from the retrieval, from May 2009 onwards, we still use two satellites for the ATOVS processing. When looking at Figs. 5 and 7 we do not see further coincidences between an apparent breakpoint and a change from two to three satellites, or vice versa. Thus, we do not expect that this breakpoint can be explained by sampling or Kriging effects. Consequently, we extended the analysis by comparing the AIRS data record to the ATOVS products for each NOAA and Metop satellite separately. During this exercise the Kriging routine is not applied to the ATOVS data. Figure 10 shows the bias between the ATOVS products derived from each NOAA and Metop satellite and the AIRS data record. During overlapping periods, data from Metop-A, NOAA-16 and NOAA-19 exhibit similar biases relative to AIRS, with values around $2\,\mathrm{kg\,m^{-2}}$. Also noticeable is the increase in bias for NOAA-16 between 2003 and 2009 and the decrease in bias after the maximum in 2009. All NOAA satellites typically have different Equator-crossing times and exhibit a drift in Equator-crossing time (see, for example, John et al., 2012, their Fig. 4). NOAA-16's orbital drift is the strongest and ranges from 14:00 local time in 2003 to 17:30 local time in 2009 and to 19:30 in 2011. The AIRS orbit is stable, with an Equator-crossing time of 13:30. Thus, at the beginning of the bias time series, the difference in temporal sampling is minimal. If the difference in Equator-crossing time were the dominant contributor to the bias, the maximum in bias can be expected at 19:30. It seems again that the diurnal cycle of deep convection in combination with differences in temporal sampling impacts the bias between AIRS and ATOVS. In addition to such a sampling error, the retrieval uncertainty might also be affected by cloud handling, which then results in a diurnal cycle of the retrieval error in the presence of convective clouds. What is most obvious is that the bias for the data derived from NOAA-15 is systematically smaller than the bias for the data derived from any of the other satellites. The average difference between the NOAA-15 and AIRS biases and between those of NOAA-16 and AIRS is almost $1\,\mathrm{kg\,m^{-2}}$. Consequently, with the removal of NOAA-15 from the retrieval, the bias relative to AIRS and GUAN data records increases. We recomputed the bias between the TPW from ATOVS and AIRS data records with NOAA-15 data being removed from the retrieval in June 2008. The results are shown in Fig. 11. The breakpoint now clearly appears in June 2008 instead of May 2009.

5 Assumptions and known limitations

The ATOVS data record was processed using a frozen processing system and up-to-date tools and retrievals. The ATOVS data record is suitable for the following applications: process studies, variability analysis, and model evaluation if the assumptions and limitations described in this section are kept in mind.
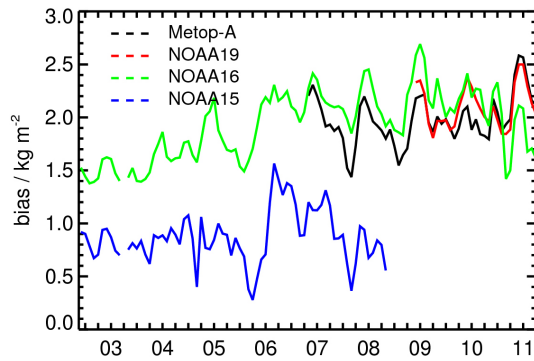
Figure 10. Bias between the ATOVS data record, derived from each satellite separately, and the AIRS V5 data record for the time period January 2003 to December 2011.
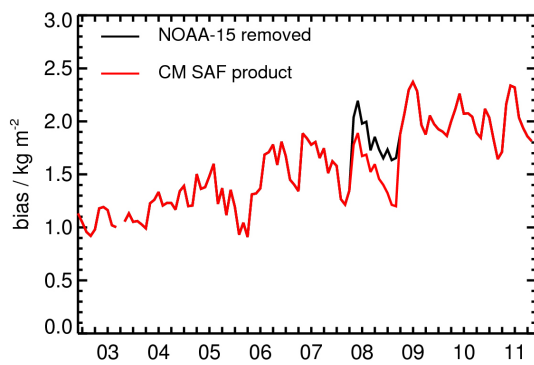

Figure 11. TPW bias between the ATOVS and the AIRS V5 data records, as in Fig. 7. Furthermore, the figure shows the bias between the ATOVS data record without the use of NOAA-15 data from June 2008 onwards (instead of May 2009).

The data record is not independent of the ERA-Interim model fields since those are used as input to the retrieval. Considering the weighting functions of the ATOVS instruments, the results in the lower troposphere over land surface may be significantly influenced by the model fields. Another related limitation is that the ERA-Interim model fields are not independent of ATOVS since the ATOVS data are assimilated in the reanalysis model.

Different satellites are used to generate the data record, and the number of satellites which are used for the processing also varies from one to four. The satellites have different local overpass times and some of them drifted with time – these two factors might affect the performance of the data record. Furthermore, the data exhibit a lower quality if only one satellite is used to generate the data record because the Kriging routine then uses morning and afternoon orbits to estimate the local variance. This is only possible if the morning and afternoon observations are valid at the same location, which reduces the number of valid observations. This impacts the quality of the first 2 years of the ATOVS data record.

The quality of the product depends on the intercalibration of the AMSU-A, AMSU-B/MHS, and HIRS brightness temperatures. A missing or nonoptimal intercalibration might lead to artifact trends. A feasible intercalibration for AMSU-A and HIRS brightness temperatures was not available at the time of processing. Only intercalibration coefficients for AMSU-B channels have been applied for the time period 2001 to 2010 (John et al., 2012). AMSU-B/MHS brightness temperatures are intercalibrated using the SNO method described in John et al. (2012). It is shown in John et al. (2012) that the measurements taken into account in the SNO occur only at the poles, and thus only a small part of the dynamic range of the global measurements is represented in the SNO. Consequently, potential non-linear effects as a function of scene brightness temperature are not considered. It has also been shown that there might be scan asymmetry in the AMSU-B brightness temperatures (Buehler et al., 2005; John et al., 2013), which has not been accounted for here.

The impacts of the Kriging and the lack of intercalibration reduce the stability of the ATOVS product.

This, in combination with missing bias correction, has a complex impact on the systematic error of the product and, together with the limited temporal coverage, makes this product unsuitable for climate change analysis.

The water vapour retrieval is not reliable in the case of very elevated terrain (mostly in the Himalaya region), because in such regions the sounders "see" through the entire atmosphere down to the surface and the signal is contaminated with surface contributions.

6 Conclusions

We introduced the recently released global CM SAF ATOVS tropospheric temperature and water vapour data record. The data record has a spatial resolution of $90\,km \times 90\,km$ and is provided as daily and monthly averages. The product suite contains TPW, LPW, layer mean temperature, and specific humidity and temperature at six pressure levels and is based on a maximum likelihood solution retrieval and postprocessing by a Kriging algorithm to allow for gap filling. The ATOVS data record covers the period January 1999 to December 2011 and has been generated by the consistent application of a fixed processing chain. The reprocessing resulted in an improvement of the quality and the stability of the ATOVS data record relative to the operational ATOVS product. To our knowledge this is the first time that an ATOVS reprocessing effort has been conducted. The ATOVS data record, in NetCDF format, and the related documents (product user manual, validation report, algorithm theoretical basis document) are freely available from the CM SAF website at http://www.cmsaf.eu/wui and http://www.cmsaf.eu/docs. The data record is referenced under doi:10.5676/EUM_SAF_CM/WVT_ATOVS/V001.

The analysis of the global TPW average from the ATOVS data record revealed a significant breakpoint between January and February 2001 which coincides with the change from the use of observations from one satellite for the retrieval to the use of two satellites. An example of the monthly mean PDF analysis shows that the Kriging systematically fills the PDF at large values. As gaps typically occur in association with precipitation, and consequently in areas of high humidity content, this is reasonable. Thus, the gap filling process through Kriging largely explains the breakpoint. We do not recommend using the ATOVS data record for the period from January 1999 to January 2001 for variability analysis due to a questionable applicability of the Kriging algorithm in the presence of data from a single satellite. Further analysis is needed to quantify the bias potentially caused by sampling gaps in the presence of precipitation, as also recommended in Schröder et al. (2013) as a result of the third G-VAP workshop.

The TPW and LPW products from the ATOVS data record have been compared to corresponding radiosonde observations at GUAN stations and to the AIRS version 5 data record in order to identify potential issues in the ATOVS data record and to characterise the ATOVS data record in a relative sense. The breakpoint between January and February 2001 is not evident in the bias between ATOVS and GUAN due to the collocation procedure. Based on the comparison to the GUAN data record, we find an averaged bias and an averaged RMSE of -0.2 and $3.3\,\mathrm{kg\,m^{-2}}$, respectively, for TPW. For LPW maximum relative (absolute) biases and RMSE are found for the highest (lowest) layer, similar to results presented in the literature. The bias between ATOVS and AIRS differs between the NOAA satellites and is minimal for NOAA-15. The next step that is needed to improve the ATOVS data record is the implementation of a bias correction scheme which needs to account for the various uncertainties of the retrieval including calibration uncertainties. The bias correction thus needs to be a function of satellite. This and other adaptations to the IAPP retrieval would require close cooperation with the University of Wisconsin, also through the International TOVS Working Group. Relative to AIRS the RMSE is typically smaller, while the bias is larger and exhibits a breakpoint between April and May 2009. At that time, NOAA-15 observations were removed from the retrieval. We further discussed the spatial distribution of the bias between the ATOVS and AIRS data records. Maximum biases coincide with regions of strong and frequent precipitation in the ITCZ – here, in particular, tropical land areas – and in the storm track regions. The bias can to a large extent be explained by (1) gap filling through Kriging, (2) clear-sky bias, (3) differences in cloud and precipitation handling, (4) differences in the handling of surface emissivities and (5) differences in diurnal sampling. Over ocean the dominating contributor is the Kriging approach. Over land a separation into the individual contributors to the overall bias is a major effort as it requires, among other things, restarting the retrievals with common input, cloud and rain screening and cloud removal. Within G-VAP, the intercomparison of gridded long-time-series satellite data records also exhibits largest discrepancies over tropical land surfaces, and further analysis will be conducted to find explanations for this. Here, we conclude that the ATOVS data record should be considered with care over tropical land surfaces and also at high elevation.

The provision of vertically resolved data in the upper troposphere is crucial for, among other things, the analysis of the water vapour feedback. In order to ease comparisons and to enhance the reliability of related conclusions, the provision of the retrieval uncertainty and averaging kernels at the pixel level would be beneficial. In the case of the gridded CM SAF product, the first of the next steps will be the implementation of the retrieval error and error propagation into the gridded product.

Finally, it became obvious that the ATOVS data record will benefit from carefully quality-controlled, recalibrated and intercalibrated sensor data. Such high-quality level 1 data are being generated in cooperation between CM SAF and EUMETSAT and within the European Union project "Fidelity and Uncertainty in Climate data records from Earth Observations".

Metop-B is the last satellite carrying the ATOVS sensor suite. The processing needs to be adapted to account for the replacement of HIRS with IASI. The quality and usability would benefit from the inclusion of data from other hyperspectral infrared and microwave sounders and from a backward extension of the processing by implementing TOVS data.

References

Andersson, A., Fennig, K., Klepp, C., Bakan, S., Graßl, H., and Schulz, J.: The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data – HOAPS-3, Earth Syst. Sci. Data, 2, 215–234, doi:10.5194/essd-2-215-2010, 2010.

Atkinson, N.: AAPP Overview, Satellite Application Facility for Numerical Weather Prediction, Document NWPSAF-MO-UD-004, Version 7.0, 2011.

August, T., Hultberg, T., Crapeau, M., and Calbet, X.: IASI Level 2 version 6. 19th International TOVS Study Conference (ITSC-19), Jeju Island, South Korea, 25 March–1 April 2014.

Bedka, S., Knuteson, R., Revercomb, H., Tobin, D., and Turner, D.: An assessment of the absolute accuracy of the Atmospheric Infrared Sounder v5 precipitable water vapor product at tropical, midlatitude, and arctic ground-truth sites: September 2002 through August 2008, J. Geophys. Res., 115, D17310, doi:10.1029/2009JD013139, 2010.

Boukabara, S.-A., Garrett, K., Chen, W., Iturbide-Sanchez, F., Grassotti, C., Kongoli, C., Chen, R., Liu, Q., Yan, B., Weng, F., Ferraro, R., Kleespies, T., and Meng, H.: MiRS: An All-Weather 1DVAR Satellite Data Assimilation & Retrieval System, IEEE Trans. Geosci. Remote Sens., 49, 3249–3272, doi:10.1109/TGRS.2011.2158438, 2011.

Buehler, S. A., Kuvatov, M., John, V. O., Leiterer, U., and Dier, H.: Comparison of microwave satellite humidity data and radiosonde profiles: A case study, J. Geophys. Res., 109, D13103, doi:10.1029/2004JD004605, 2004.

Buehler, S. A., Kuvatov, M., and John, V. O.: Scan asymmetries in AMSU-B data, Geophys. Res. Lett., 32, L24810, doi:10.1029/2005GL024747, 2005.

Christy, J. R. and Norris, W. B.: Discontinuity Issues with Radiosonde and Satellite Temperatures in the Australian Region 1979–2006, J. Atmos. Oceanic Technol., 26, 508–522, doi:10.1175/2008JTECHA1126.1, 2009.

Courcoux, N. and Schröder, M.: ATOVS tropospheric humidity and temperature data set: Algorithm Theoretical Basis Document – ATOVS edition 1. EUMETSAT CM SAF ATBD, reference number: SAF/CM/DWD/ATOVS, revision 1.1, 13 March 2013.

Daan, H.: Guide to the GCOS surface and upper air networks: GSN and GUAN, Version 1.1, GCOS-73, WMO/TD No. 1106, 2002.

Dai, A., Wang, J., Ware, R. H., and Van Hove, T. M.: Diurnal variation in atmospheric water vapor over North America and its implications for sampling errors in radiosonde humidity. J. Geophys. Res., 107, CL11-1–ACL11-14, doi:10.1029/2001JD000642, 2002.

Dai, A., Wang, J., Thorne, P. W., Parker, D. E., Haimberger, L., and Wang, X. L.: A new approach to homogenize daily radiosonde humidity data, J. Climate, 24, 965–991, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the

data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, 2011.

Dessler, A. E. and Sherwood, S. C.: A Matter of Humidity, Science, 323, 1020–1021, doi:10.1126/science.1171264, 2009.

Divakarla, M., Barnet, C., Liu, X., Gu, D., Wilson, M., Kizer, S., Xiong, X., Maddy, E., Ferraro, R., Knuteson, R., Hagan, D., Ma, X., Tan, C., Nalli, N., Reale, A., Mollner, A. K., Yang, W., Gambacorta, A., Feltz, M., Iturbide-Sanchez, F., Sun, B., and Goldberg, M.: The CrIMSS EDR Algorithm: Characterization, Optimization, and Validation, J. Geophys. Res. Atmos., 119, 4953–4977, doi:10.1002/2013JD020438, 2014.

Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, J. Climate, 19, 53–68, 2006.

Eyre, J. R. and Hilton, F. I.: Sensitivity of analysis error covariance to the mis-specification of background error covariance, Q. J. Roy. Meteorol. Soc., 139, 524–533, doi:10.1002/qj.1979, 2013.

Gambacorta, A., Barnet, C., Wolf, W., Goldberg, M., King, T., Nally, N., Xiong, X., Maddy, E., and Divakarla, M.: The NOAA Unique CrIS/ATMS processing system (NUCAPS): First light results, paper presented at Proceedings of the International TOVS Working Group Meeting, ITSC-XVIII, Toulouse, France, 2012.

GCOS-82: The second report on the adequacy of the global observing systems for climate in support of UNFCCC, GCOS–82, WMO/TD No. 1143, World Meteorological Organization, Geneva, Switzerland, available at: www.wmo.int/pages/prog/gcos (last access: 11 November 2015), April 2003.

Held, I. M. and Soden, B. J.: Water vapor feedback and global warming, Annu. Rev. Energy Environ., 25, 441–475, 2000.

Hurst, D. F., Oltmans, S. J., Vömel, H., Rosenlof, K. H., Davis, S. M., Ray, E. A., Hall, E. G., and Jordan, A. F.: Stratospheric water vapor trends over Boulder, Colorado: Analysis of the 30 year Boulder record, J. Geophys. Res., 116, D02306, doi:10.1029/2010JD015065, 2011.

John, V. O., Holl, G., Buehler, S. A., Candy, B., Saunders, R. W., and Parker, D. E.: Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses, J. Geophys. Res., 117, D02305, doi:10.1029/2011JD016349, 2012.

John, V. O., Holl, G., Atkinson, N., and Buehler, S. A.: Monitoring scan asymmetry of microwave humidity sounding channels using simultaneous all angle collocations (SAACs), J. Geophys. Res., 118, 1536–1545, doi:10.1002/jgrd.50154, 2013.

Kiehl, J. T. and Trenberth, K. E.: Earth's annual global mean energy budget, Bull. Amer. Meteor. Soc., 78, 197–208, 1997.

Kleespies, T. J. and Watts, P.: Comparison of simulated radiances, jacobians and linear error analysis for the Microwave Humidity Sounder and the Advanced Microwave Sounding Unit–B, Q. J. Roy. Meteorol. Soc., 132, 3001–3010, 2006.

Kunz, A., Müller, R., Homonnai, V., Jánosi, I., Hurst, D., Rap, A., Forster, P., Rohrer, F., Spelten, N., and Riese, M.: Extending water vapor trend observations over Boulder into the tropopause region: trend uncertainties and resulting radiative forcing, J. Geophys. Res., 118, 11269–11284, doi:10.1002/jgrd.50831, 2013.

Labrot, T., Lavanant, L., Whyte, K., Atkinson, N., and Brunel, P.: AAPP documentation scientific description, Satellite Application Facility for Numerical Weather Prediction, Document NWPSAF-MF-UD-001, Version 7.0, 2011.

Labrot, T., Aktinson, N., and Roquet, P.: AAPP documentation software description, Satellite Application Facility for Numerical Weather Prediction, Document NWPSAF-MF-UD-002, Version 7.1, 2012.

Leiterer, U., Dier, H., Nagel, D., Naebert, T., Althausen, D., Franke, K., Kats, A., and Wagner, F.: Correction method for RS80-A Humicap humidity profiles and their validation by Lidar backscattering profiles in tropical cirrus clouds, J. Atmos. Oceanic Tech., 22, 18–29, 2005.

Li, J., Wolf, W. W., Menzel, W. P., Zhang, W., Huang, H.-L., and Achtor, T. H.: Global sounding of the atmosphere from ATOVS measurements: the algorithm and validation, J. Appl. Meteorol., 39, 1248–1268, 2000.

Mieruch, S., Schröder, M., Noel, S., and Schulz, J.: Comparison of monthly means of global total column water vapor retrieved from independent satellite observations, J. Geophys. Res., 115, D23310, doi:10.1029/2010JD013946, 2010.

Miloshevich, L., Paukkunen, M., Vömel, H., and Oltmans, S. J.: Development and validation of a time-lag correction for Vaisala radiosondes humidity measurements, J. Atmos. Ocean. Tech., 21, 1305–1327, 2004.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., Lesht, B. M., Schmidlin, F. J., and Russo, F.: Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEXG and implications for AIRS validation, J. Geophys. Res., 111, D09S10, doi:10.1029/2005JD006083, 2006.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., and Leblanc, T.: Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements, J. Geophys. Res., 114, D11305, doi:10.1029/2008JD011565, 2009.

Moradi, I., Buehler, S. A., and John, V. O.: Comparing upper tropospheric humidity from microwave satellite instrument and IGRA radiosonde data, MicroRad processing, 978-1-4244-8121-7, 2010.

Pougatchev, N., August, T., Calbet, X., Hultberg, T., Oduleye, O., Schlüssel, P., Stiller, B., Germain, K. St., and Bingham, G.: IASI temperature and water vapor retrievals – error assessment and validation, Atmos. Chem. Phys., 9, 6453–6458, doi:10.5194/acp-9-6453-2009, 2009.

Randel, D. L., Von der Haar, T. H., Ringerud, M. A., Stephens, G. L., Greenwald, T. J., and Combs, C. L.: A new global water vapor data set, Bull. Amer. Meteor. Soc., 77, 1233–1246, 1996.

Reale, A., Tilley, F., Ferguson, M., and Allegrino, A.: NOAA operational sounding products for ATOVS, Int. J. Remote. Sens., 29, 4615–4651, 2008.

Reale, T., Sun, B., Tilley, F. H., and Pettey, M.: The NOAA Products Validation System (NPROVS), J. Atmos. Ocean. Tech., 29, 629–645, doi:10.1175/JTECH-D-11-00072.1, 2012.

Schneider, N., Schröder, M., Steenbergen, T., and Stengel, M.: ESA DUE GLOBVAPOUR Product Validation Report (PVR) for SSM/I+MERIS, issue 3, revision 1, 21 August 2012.

Schröder, M. and Lockhoff, M.: Intercomparison of trend results. Presented at the 3rd G-VAP workshop at CIRA, CSU, CO, USA in October 2013, available at: http://gewex-vap.org/?page_id= 529 (last access: 11 November 2015), 2013.

Schröder, M., Jonas, M., Lindau, R., Schulz, J., and Fennig, K.: The CM SAF SSM/I-based total column water vapour climate data record: methods and evaluation against re-analyses and satellite,

Atmos. Meas. Tech., 6, 765–775, doi:10.5194/amt-6-765-2013, 2013a.

Schröder, M., Lockhoff, M., and Shi, L.: 3rd GEWEX Water Vapor Assessment (G-VAP) Workshop, GEWEX Newsletter, 23, November 2013b.

Scott, N. A., Chédin, A., Armante, R., Francis, J., Stubenrauch, C., Chaboureau, J. P., Chevallier, F., Claud, C., and Chéruy, F.: Characteristics of the TOVS Pathfinder Path-B data set, Bull. Amer. Meteor. Soc., 80, 2679–2701, 1999.

Shi, L. and Bates, J. J.: Three decades of intersatellite-calibrated High-Resolution Infrared Radiation Sounder upper tropospheric water vapour, J. Geophys. Res., 116, D04108, doi:10.1029/2010JD014847, 2011.

Soden, B. J. and Lanzante, J. R.: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor, J. Climate, 9, 1235–1250, 1996.

Sohn, B. J. and Bennartz, R.: Contribution of water vapor to observational estimates of longwave cloud radiative forcing, J. Geophys. Res., 113, D20107, doi:10.1029/2008JD010053, 2008.

Sun, B., Reale, A., Seidel, D. J., and Hunt, D. C.: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among types and the effects of imperfect collocation on comparison statistics, J. Geophys. Res., 115, D23104, doi:10.1029/2010JD014457, 2010.

Susskind, J., Barnet, C. D., and Blaisdell, J. M.: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, IEEE Trans. Geosci. Remote Sens., 41, 390–409, 2003.

Susskind, J., Barnet, C., Blaisdell, J., Iredell, L., Keita, F., Kouvaris, L., Molnar, G., and Chahine, M.: Accuracy of geophysical parameters derived from AIRS/AMSU as a function of fractional cloud cover, J. Geophys. Res., 111, D09S17, doi:10.1029/2005JD006272, 2006.

Susskind, J., Blaisdell, J., Iredell, L., and Keita, F.: Improved Temperature Sounding and Quality Control Methodology Using AIRS/AMSU Data: The AIRS Science Team Version 5 Retrieval Algorithm IEEE Trans. Geosci. Remote Sens., 49, 883–907, doi:10.1109/TGRS.2010.2070508, 2011.

Trenberth, K. and Stepaniak, D.: Covariability of components of poleward atmospheric energy transports on seasonal and interannual timescales, J. Climate, 16, 3691–3705, 2003a.

Trenberth, K. and Stepaniak, D.: Seamless Poleward Atmospheric Energy Transports and Implications for the Hadley Circulation, J. Climate, 16, 3706–3722, 2003b.

Vömel, H., Selkirk, H., Miloshevich, L., Valverde-Canossa, J., Valdés, J., Kyrö, E., Kivi, R., Stolz, W., Peng, G., and Diaz, J. A.: Radiation dry bias of the Vaisala RS92 humidity sensor, J. Atmos. Ocean Tech., 24, 953–963, doi:10.1175/JTECH2019.1, 2006.

Vonder Haar, T. H., Bytheway, J. L., and Forsythe, J. M.: Weather and climate analyses using improved global water vapor observations, Geophys. Res. Lett., 39, L15802, doi:10.1029/2012GL052094, 2012.

Waller, J. A., Dance, S. L., Lawless, A. S., Nichols, N. K., and Eyre, J. R.: Representativity error for temperature and humidity using the MetOffice high-resolution model, Q. J. Roy. Meteorol. Soc., 140, 1189–1197, doi:10.1002/qj.2207, 2014.

www.earth-syst-sci-data.net/7/397/2015/

Earth Syst. Sci. Data, 7, 397–414, 2015

Yang, G.-Y. and Slingo, J.: The Diurnal Cycle in the Tropics, Mon. Weather Rev., 129, 784–801, doi:10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2, 2001.

Zou, C.-Z. and Wang, W.: Intersatellite calibration of AMSU-A observations for weather and climate applications, J. Geophys. Res., 116, D23113, doi:10.1029/2011JD016205, 2011.