**Interactive comment on “The ESA GOME-Evolution “Climate” water vapor product: A homogenized time-series of H$_2$O columns from GOME/SCIAMACHY/GOME-2” by Steffen Beirle et al.**

R. Lang (Referee)
ruediger.lang@eumetsat.int

Received and published: 22 November 2017

Water vapour retrievals from the visible region (here the red part) of the spectrum contribute to the overall suite of available water vapour remote sensing retrievals by offering a very good accuracy over land and a very acceptable accuracy over ocean under day-light conditions. Like their companion retrievals from the thermal infrared region, VIS-NIR WV retrievals are sensitive to clouds, which have to be screened out, at least to a large extent. The suite of available retrieval methods for UV-VIS-NIR instruments like GOME-1 and 2, and SCIAMACHY (and also OMI) have also the advantage
to be relatively free of a priori knowledge of the atmospheric state and can therefore be considered as truly independent measurements both with respect to ground-based measurements and WV concentrations provided by model output.

The provision of a quality controlled and consistently reprocessed merged data-set from the GOME-1 and 2 and SCIAMACHY suite of instruments, already covering two decades of data and (with Metop-B and C) to be continued until the end of the 2020th, therefore provides a very important basis for validation and trend analysis.

The paper by Beirle et al. describes the strategies used to merge the three WV data-sets available using a retrieval strategy originally developed by Wagner et al and implemented in the GOME operational processor environment (GDP) by DLR (Grossi et al.). Since the retrieval method, using the 640 nm region, as well as the conversion from SCD to VCDs using AMFs derived from the oxygen B-band, are the same for all instruments, the only remaining issue is to overcome instrument design and operation specific differences and their impact on the climatological WV values derived as “cloud-free” monthly means at global coverage. Cloud screening is carried out by comparing the retrieved ground-pixel averaged mean scattering-height (derived from the O2-AMF) to the theoretical cloud-free height (using prescribed “maximum” pressure profiles per instrument above a cloud-free pacific-ocean), while the strategy for merging of the data, overcoming instrument specific design aspects, is to reduce the spatial resolution of GOME-2 and SCIAMACHY to that of the approximately 8 times coarser GOME-1 foot-print and subtract a global (SCIAMACHY) or latitude depended (GOME-2) offset.

The paper is well written and the results are scientifically sound so that I can recommend it for publication in ESS, provided the authors can address the two main issues I have with the presented results and analysis.

I) The reason why the offsets between SCIA and GOME-2 are so different for land and ocean is still very puzzling after having read the paper, since for me it seems
that this cannot be explained by the interference of the missing/screened out monthly narrow scan mode (NSM). This is because such instrument operations issues, like most calibration issues, should be largely independent from the surface type, as also Figure A1 - showing the impact of the NSM removal on the mean scan angle - seems to demonstrate. One can only guess that the observed land/ocean difference in the offset between the two instruments WV data might be more related to the SCD retrievals, e.g. in case they use different surface treatments or databases or similar auxiliary data related to different surface types.

II) Statistical effects from monthly averages in combination with cloud screening are not considered. This is an oversight which is often made when evaluating monthly mean data-series, especially for species with such high temporal variability. If the cloud screening per instrument, and during the course of one month, is different due to differences in spatial sampling or due to the different coverage (e.g. by the SCIAMACHY alternate limb/nadir sounding), this can lead to significant monthly differences in the averaged result, simply because the monthly average covers different days, and since the day-to-day WV variations at the same point can be very large. This may in particular play a significant role for the gaps analysis (Section 4.2), and in particular in the tropics where the temporal variability (in absolute terms) is very high. So a comparison of the WV distribution around the mean in selected months (or for each month for a selected year) should be carried out, for analysing this effect on the presented monthly mean values.

Specific comments:

Section 2.4, p.4, last paragraph: But this “consistency” could also be achieved by, e.g., using the same independent pixel approach (e.g. FRESCO+) applied to the O2-B band. And it is not clear to me why a DEM is not used for clearly separating a mountain or a cloud (using the O2-AMF approach)?

Figure 2: Why is the corresponding GOME-2 situation not shown here? I think this
would be very useful for the reader and for the user of the data-set as well.

Editorials:

Section 1, p.2, l.7: I would add "ground-based" here since GPS can also be satellite.

Section 1, p.2, l.9+10: Radio occultation (upper troposphere/stratosphere) - should be mentioned here, since RO is meanwhile a key-contributor to NWP for WV.

Section 1, p.2, l.15: I would add here that therefore these measurements are “complementary to the MW, RO and IR derived climate data-sets, which are sensitive only to specific surfaces or altitude ranges”.

Section 1, p.2, l.19: Isn’t GDP in the version 4.8, meanwhile?

Section 1, p.2, l.29: I would indicate here already what the difference in the footprint is, by adding "to the smaller swath with larger ground footprint" ... or similar.

Section 1, p.2, l.31: I would add that this is why this makes it also a truly independent data-set for model evaluation and evaluation of other WV climatologies.

Section 2.2, p.3, l.23-24: Revise for better reading.

General: “saturation effect” should be explained somewhere, for the non-retrieval DOAS expert as being another name for the “non-linearity” effect in the spectral absorption.

Section 2.3, p.4, l.5: “by” Wang et al.

Figure 4: The caption misses a description of the different SCIAMACHY data-sets displayed.

References: Danielczok, A. and Schröder, M. - The reference is not complete here, although provided in the text.

Figure A1: Caption should state that it is a “monthly” mean.