Answer to Referee #2

L. Lelli (luca@iup.physik.uni-bremen.de)

April 28, 2014

Subject: acp-2013-821
Date: Thursday, April 24, 2014, 10:27:43 AM
From: Svenja Lange <svenja.lange@copernicus.org>
To: luca@iup.physik.uni-bremen.de
CC: tim.garrett@utah.edu

Dear Luca Lelli,

please find below another referee comment from referee 2:

"When I meant the title is misleading, I would suggest to modify the title to: 'Trends in top height of opaque clouds from passive observations ... ', to point out that only a sub-sample of clouds is studied (which is ok); the authors replied that this is written in the abstract, but it would make the title more consistent with the analysis.

In addition, the authors use further sub-sampling by discarding certain cases (depending on a quality flag). I strongly recommend showing the fraction of rejections or the fraction of clouds kept for the analysis out of all detected clouds as function of time, to study variations not only in cloud height of the sub-sample but also to see if the studied sub-sample (of homogeneous opaque clouds) stays constant."

Kind regards,
Svenja Lange

(a) Title

The optimal title would also contain informations on the time span covered by the analysis, as well as on the type of sensor used (i.e., spectrometers) because also imagers are equipped with the oxygen A-band (e.g. MERIS). However, all these informations are provided to the reader right after the title, within the first few lines of the abstract, as it should be. For these reasons, we think that the title is well balanced.

(b) Sample size

If we correctly guess the concern of the referee, the issue is whether the time series of retrievals (i.e., counts) used for the analysis is stable. Other said, if the studied sub-sample is populated enough to provide statistical robustness and/or if the algorithm performance degrades in time.

Table 1: SACURA quality flags.

Value	Description
0	No retrieval
1	Only cloud bottom height convergence
2	Only cloud top height convergence
3	Geometrical thickness limit
4	No convergence
5	Cloud top and bottom height convergence

To this end, many aspects deserve attention.

First off, please note that in view of Eq. (2) in the paper, the monthly-sampled columnar CTH is already weighted with the relative local (i.e., for each height bin) retrieval counts. This is because spatial resolution differs among the instruments (see below). In this way, possible inhomogeneities in the CTH sample size are already taken into account.

Second, it is worth stressing that a cloud retrieval algorithm can't suffer from lack of statistical representativeness due to limited number of observations, since clouds cover on average $\approx 65\%$ of the globe, at any given time and location. For instance, this is not true for remote sensing of aerosols, for which cloud clearance is of uppermost importance. What is regarded as *disturbance* in the trace gases and aerosol communities, here is the *signal*.

Third, the oxygen A-band is located in the NIR. Past experience has shown that the spectrometers used in this work do not almost suffer from degradation in this spectral range, whereas UV bands are more prone to degradation. More important, what matters in the exploitation of the A-band is the ratio between the line core around 761 nm and the continuum at 758 nm. Holding cloud properties constant within a ground pixel, this ratio stays quite stable in time, making the A-band a "self-calibrating" spectral window. For these reasons the SACURA algorithm, due to its design, benefits from the inherent stability of the A-band.

Obviously, the time series of observations have been investigated prior to preparation of the paper. Figure 1 shows the occurrences of the quality flags (see Table 1 for their meaning) for the complete dataset of the three instruments as function of cloud fraction (left column), cloud optical thickness (mid column) and cloud top height (right column). The corresponding statistics are given in Table 2. Table 2 shows also that the number of usable retrievals increases and the relative number of rejections drops as function of instrument. In Figure 2 the time series of the relative sample sizes used for the derivation of the anomalies (Fig. 11 of the manuscript) are portrayed. It can be seen that the sub-sample is stable, as function of time, for each instrument.

We think that the discussion above stems from well-known facts and, if added, would divert the reader from the current flow of the paper. Therefore, no change will be introduced in the manuscript.

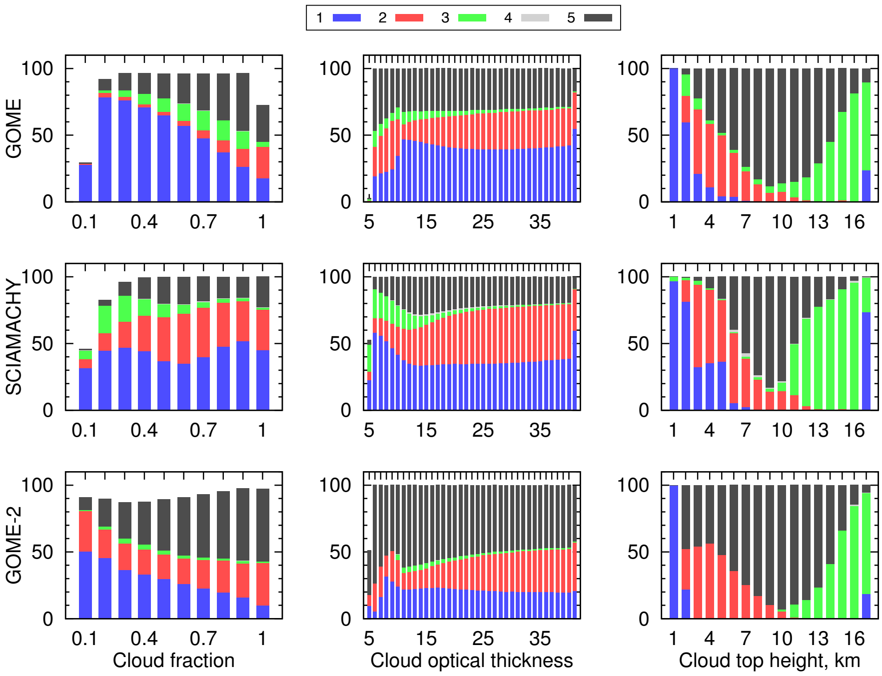

Figure 1: Quality flag statistics of cloud property retrievals used for the analysis, normalized to the total number of retrievals. Note that missing values in the histograms stand for quality flag 0 (“no retrieval”) and aren’t written to the output. The studied sub-sample is populated with black, red and green retrievals, given a CTH $<$ 5 km.

Table 2: Statistics of the quality flags for GOME (total number of ground pixels 41,183,749), SCIAMACHY (204,406,630), and GOME-2 (95,208,916). Values are given in % of the total number.

Flag	GOME	SCIA	GOME-2
0	23.88	8.11	6.11
1	34.6	42.47	20.51
2	13.71	25.94	25.15
3	5.69	8.19	2.11
4	0.02	0.40	0.04
5	22.10	14.89	46.1

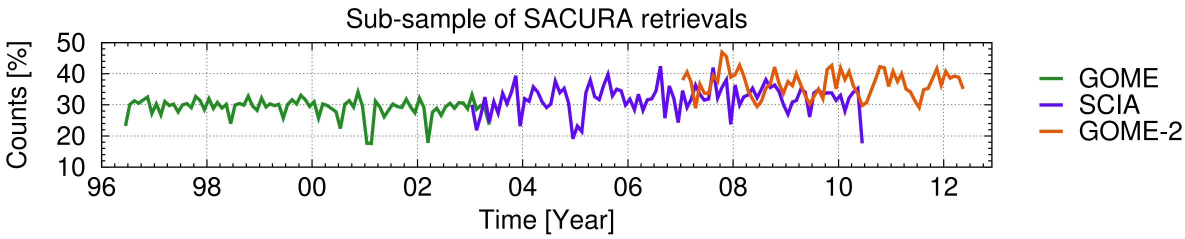

Figure 2: Time series of relative counts in the latitude belt ±60° used for the calculation of the anomaly time series.