Author response
Luca Lelli (luca@iup.physik.uni-bremen.de)
April 30, 2014

This document has the following structure:

Reviewer comment
Author response
(Action) changes introduced in the manuscript

Answer to Anonymous Referee #1

Specific comments:
1) p. 31410 Lines 23-25: need better explanation of how high cloud absorbing and radiating warms the atmosphere (hint: high clouds are cold).
2) p. 31411 lines 1-6: The sentence structure is awkward.

The following rewording is proposed:

"The altitude of a cloud plays a cardinal role because clouds at different heights (i.e., temperatures) exert different feedbacks. High-altitude clouds absorb infrared radiation that comes from the lower atmosphere and radiate like blackbodies. Since the temperature contrast between elevated clouds and lower atmosphere is high, the clouds have a positive contribution on the local net energy balance. At the same time, their albedo is small because they are, in most cases, comparably thin. Thus, they can warm the atmosphere more than they cool it, exerting a positive feedback.

Conversely, low-altitude clouds are strongly reflecting objects owing to their high optical density but they ineffectually shield infrared radiation emitted by atmospheric gases to the outer space. The reason is that the temperature of low clouds is closer to the temperature of the ground and the local net energy balance is close to zero. These clouds in turn can cool the climate system more than they warm it and thus exert a negative feedback.

Loeb et al., 2012 observed this mechanism in a study that focused on clouds in the tropical belt. However, this situation may change when considering the Pacific Northeast over a decadal time window. Evidence of a positive feedback by low-level clouds has already been demonstrated (Clement et al., 2009). It is therefore likely that no general description is possible on a global scale and regional studies should be conducted instead."

Action (Sect. 1, l. 19–35) The cloud feedback mechanism as function of cloud altitude has been reworded as proposed above.

3) p. 31412: It would be helpful to have some more detail on how cloud fraction and cloud reflectance are determined since there can be trade-offs between the two in Eq. 1. E.g., measured scene reflectance can result from either/both greater cloud fraction and greater cloud reflectance, but they may have different impacts on retrieved cloud top height.

The cloud fraction is determined analyzing PMD radiances. PMDs offer a better spatial resolution within the instrument footprint and a broader spectral coverage (from UV to NIR).

Once the fractional cloud cover is determined, its value is used to scale the actual measurement within the Independent Pixel Approximation. This means that the cloud reflectance ingested in SACURA corresponds always to cloud fraction 1.

It has been demonstrated ([Kokhanovsky et al., 2007]) that the cloud reflection function it is not affected very much by horizontal photon transport, as long as cloud fraction is known from an independent source. This statement is true regardless of cloud fraction and of instrumental spatial resolution, because the algorithm makes use of spectra ratios, i.e. $(R_{758} - R_\lambda)/R_{758}$.

This effect has been demonstrated by [Lelli et al., 2012, Fig. 9, p. 1559]. The absent correlation between CF bias (defined as the difference between OCRA CF and co-located CF derived from a finer resolved instrument ATSR-2 within a GOME ground pixel) and CTH bias indirectly validates Eq. 1 of the present manuscript.

From a physical point of view, it is also clear that, on such coarse footprint scales, the TOA cloud reflectance is dominated by the contribution of photons scattered directly back to the platform. This is not true for instruments such as MODIS or (A)ATSR, for which a cloud volume may have side lengths comparable with the footprint size.

Action (Sect. 2, l. 80–88) The following paragraph has been added:

“It has been demonstrated ([Kokhanovsky et al., 2007, Lelli et al., 2012]) that cloud top height is not substantially affected by the scaling of the cloud reflection function with different values of fractional cloud cover as long as f is known from an independent source. This statement is true regardless of cloud fraction and of instrumental spatial resolution, because the algorithm makes use of spectral ratios. From a physical point of view, it is also clear that, on the coarse footprint scales of the sensors used in this work, the cloud reflectance is dominated by the contribution of photons scattered directly back to the platform and not by horizontal photon transport. This is not true for spatially better resolved instruments for which a cloud volume may have side lengths comparable with the footprint size.”

4) p. 31414 lines 5-6: Considering that getting the right correction and harmonization are critical factors for the time series analysis, more detail should be provided on how this is accomplished.

We emphasize that the time series of absolute values have been neither corrected for the impact of spatial resolution nor fitted directly. This is because the three instruments have different sensing local times. The chosen strategy has been to regress time series of anomalies, instead. This approach is analogous to the customary technique described in [Mieruch et al., 2008, Eq. 1, p. 495, and references therein]. Eq. 1 is reported below for convenience.

Being Y_t the monthly mean of the variable of interest (in this context CTH) at time t (for each geolocation point on the map), $\mu\, C_t$ the offsets of the regression line, ω the desired change rate of CTH at time step X_t and N_t the noise, in the r.h.s. of the following equation

$$Y_t = \mu\, C_t + S_t + \omega\, X_t + \delta\, U_t + N_t \tag{1}$$

the term $\delta\, U_t$ describes the level shift δ in CTH, allowed when concatenating time series of GOME and SCIAMACHY at time T_0, with a step function U_t defined as

$$U_t = \begin{cases} 0, & t < T_0 \\ 1, & t \geqslant T_0 \end{cases} \tag{2}$$

Likewise, the removal of the sample CTH mean for each respective month from the time serie of absolute values allows the seasonality (the S_t in the r.h.s. of Eq. 1) to be accounted for. The term $\delta\, U_t$ is incorporated by performing this step separately for each instrument. This is because the sample mean of anomalies is centered, by definition, about zero [Wilks 2011, Sect. 3.4.2, 4.4.2] and the constant μ can be neglected (being μ the mean water vapour column of the time series at time t=0 [Mieruch et al., 2008]). Eventually, Eq. 1 reduces to Eq. 3 of the manuscript and potential autocorrelative effects are embedded in the noise term N_t (i.e., ϵ_t in our paper).

Action (Sect. 3, l. 164–182) The following paragraphs have been added.

"We emphasize that the time series of absolute values have been neither corrected for the impact of spatial resolution nor fitted directly. This is because the time series of Fig. 1 are shifted and the three instruments have different sensing local times. The chosen strategy has been to regress time series of anomalies, instead. This approach is analogous to the customary technique described in [Mieruch et al., 2008, Eq. 1, p. 495, and references therein] and is reported below for convenience.

Being Y_t the monthly mean of the variable of interest at time t (for each geolocation point on the map), $\mu\, C_t$ the offsets of the regression line, S_t the seasonal component, ω the desired change rate of the variable at time step X_t and N_t the noise, in the r.h.s. of the following equation

$$Y_t = \mu\, C_t + S_t + \omega\, X_t + \delta\, U_t + N_t$$

the term $\delta\, U_t$ describes the level shift δ allowed when concatenating time series from different instruments at time T_0, with a step function U_t defined as

$$U_t = \begin{cases} 0, & t < T_0 \\ 1, & t \geqslant T_0 \end{cases}$$

The removal of the sample mean for each respective month from the time serie of absolute values allows the seasonality to be accounted for. The term $\delta\, U_t$ is incorporated by performing this step separately for each instrument. This is because the sample mean of anomalies is, by definition, centered about zero [Wilks 2011] and the constant μ can be neglected (being μ the mean value of Y_t at time t=0). Eventually, Eq. 4 reduces to Eq. 3 and potential autocorrelative effects are embedded in the noise term."

5) p. 31414 lines 15-16: I would have greater confidence in the harmonization of GOME and SCIAMACHY time series if there were more overlap between the two and they agreed well. There is more overlap between SCIAMACHY and GOME-2, but they do not agree well, which does not give me confidence that the authors can reliably detect changes in cloud top height. Joining at June 2008 seems arbitrary, and the authors do not offer any justification for that point.

First, GOME time serie is limited to May 2003 due to a failure of the on-board tape recorder and global coverage wasn't provided from June 2003 onward. SCIAMACHY didn't reach final flight conditions before January 2003. So, only five months could be matched between GOME and SCIAMACHY.

Second, June 2008 has been chosen because it is not only the time when GOME-2 and SCIAMACHY time series converged, but also because when the PMD band definitions have been upgraded. Given that SCIAMACHY is a well calibrated instrument, we used its monthly means prior 06/2008. For the rest of the time serie, both instruments sense the dip at 2011. This feature dominates both order of magnitude and sign of the trend (the main focus of the paper), which would be preserved with SCIAMACHY monthly means anyway.

For the future, we are confident that a new reprocessing (together with a longer time coverage provided by adding MetOp-B retrievals to the time serie) will improve the significance of this kind of analysis.

Action: (Sect. 2, l. 125–145) The following paragraphs have been rewritten.

"While the CTH anomalies from GOME and SCIAMACHY are in almost perfect agreement, the transition between SCIAMACHY and GOME-2 records is not smooth. The reason for discrepancies may be manifold. On one hand, a change in the PMD pixel definitions of GOME-2 was devised in April 2008 (EUMETSAT, 2010), which may have impacted cloud fraction f and the cloud reflection function R_{cl} of Eq. 1. On the other hand, radiometric calibration issues may also influence retrievals inferred from different instruments. For instance, van Diedenhoven et al. (2005) have shown that a difference of 20 hPa in surface pressure retrieved in the oxygen A-band by GOME and SCIAMACHY could be corrected adding an offset of 0.86% to the TOA reflectance in the continuum at 756 nm. However, tests have shown that radiometric uncertainties have almost no impact in the retrieved CTH with SACURA. This is a feature of the algorithm, which is based on spectral ratios and on the concurrent fit of CTH and CBH along the whole band.

For the ensuing analysis, the time series of GOME and SCIAMACHY have been joined in June 2003 because the on-board tape recorder of GOME broke down and global coverage wasn't provided after May 2003. Time series of SCIAMACHY and GOME-2 have been joined in May 2008. Given that SCIAMACHY is a thoroughly tested instrument and GOME-2 PMD definitions were updated in April 2008, we used SCIAMACHY anomalies prior May 2008, even though GOME-2 data were available. For the rest of the time serie, GOME-2 anomalies were used, as both instruments converge and exhibit good overlap. In fact, both instruments sense the strong negative CTH anomaly in year 2011. This feature dominates the whole time serie and it gives us reasonable confidence that the order of magnitude as well as the sign of the CTH trend aren't substantially affected, when using SCIAMACHY data."

6) Fig. 7: It seems like there might be too much smoothing and some shorter time scale information is lost.
7) p. 31417 lines 2-5: It's not clear what this sentence means.

First, the plot has been redone applying a running mean filter of 6 months (Fig. 1) and oscillations on shorter time scale have emerged. Second, the sentence means that the 2-month constant lag between the column-averaged CF anomaly time serie (top plot, purple curve) and the SST anomaly can be explained by the constant lag between low-level and high-level clouds. This effect is even more evident about the extrema of the respective curves in the bottom plot of Fig. 1. We will replace the figure and reword the sentence accordingly.

(Action, Fig. 8, page 22) The figure has been replaced, after less smoothing has been applied to the time series. The following sentence has been rephrased **(l. 235-237, page 8)**:
"In Fig. 8 it is also shown that the lagged extrema in column-averaged CF of Fig. 7 originate in the lowest layers of the troposphere, because CF for low- and mid-level clouds always anticipate the extrema for high clouds by a constant lag of 2 months."

8) To what extent does low/mid cloud fraction increase when high cloud fraction decreases because low/mid clouds are no longer overlapped by higher clouds? Perhaps there is no real increase in low/mid cloud fraction, but instead the clouds can be seen due to a reduction of high cloud obscuration?

This mechanism is surely true and can't be excluded, especially with such coarse footprints.

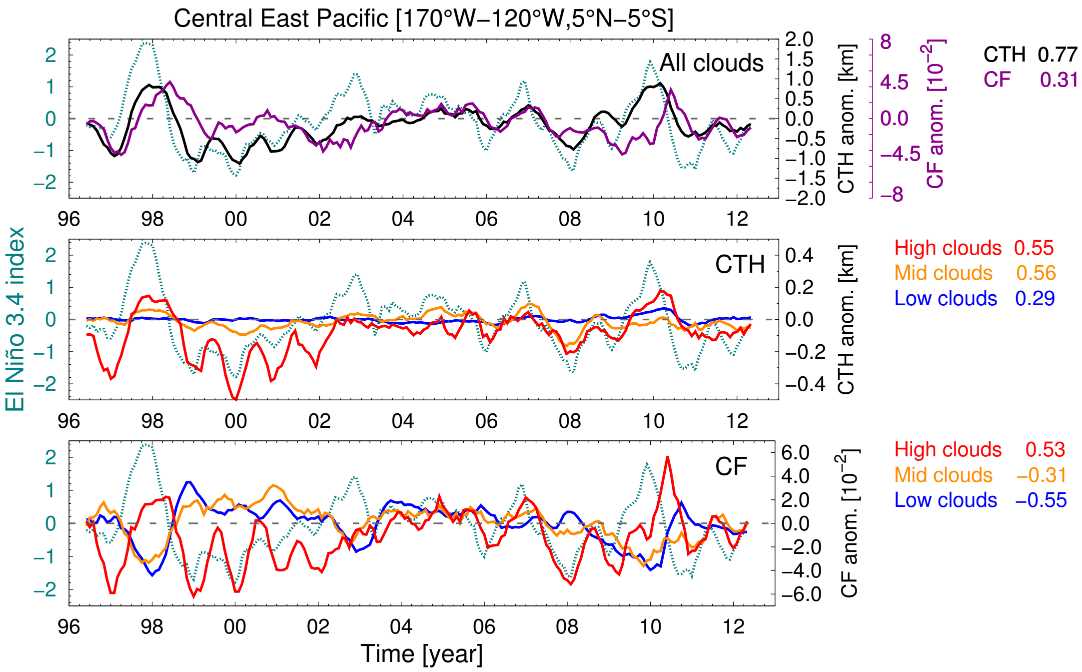

Figure 1: As in Fig. 7, p. 31438 of the manuscript, but with a running mean filter of 6 months. Time series of CTH (black curve) and CF (purple) anomaly and subsets of HC, MC and LC. All time series with their respective correlation with El Niño 3.4 index.

Being aware of this, in the conclusions (ll. 16-18 p. 31421) we worded as follows: "In particular, it has been seen that the instrumental spatial resolution impacts the calculation of mean values of apparent cloud top height at a monthly sampling", where the word "apparent" implies the inherent limitations of the instruments.

(Action) No change.

9) Fig. 8: Is this for the same region used in Figs. 6-7?

Yes.

(Action) The information has been now added in the caption.

10) Fig. 8: I can understand how cloud fraction in a particular vertical interval can change and therefore have a correlation with Nino 3.4, but how can cloud top height in a particular vertical interval change? Isn't the vertical interval fixed? Or is average height within that interval changing?

The vertical interval is fixed. What is changing within the vertical interval is the count number of CTH retrievals per height bin k (i.e., c_k at numerator of Eq. 2 in the manuscript). This can be understood as clouds being reclassified in adjacent height bins.

(Action) No change.

11) Fig. 13: Probably what matters for high clouds is not columnar water vapor, which will be

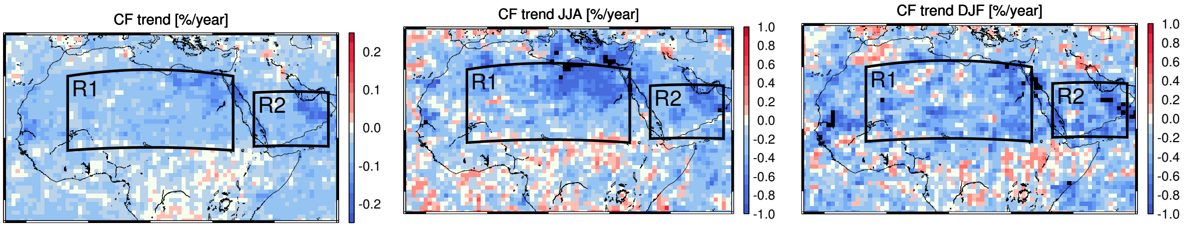

Figure 2: Trend [$\% \cdot \mathrm{year}^{-1}$] map over North Africa (R1) and Arabian Peninsula (R2) for annual cloud fraction (left plot), summer (mid) and winter (right) months. Average cloud fraction over R1 and R2 ≈ 0.4.

dominated by the boundary layer, but upper tropospheric humidity. One reason larger trends may be seen over North Africa and Arabia is that there are few low-level clouds in those regions due to desert conditions. Thus any increase in high-level clouds will be less diluted by the presence of low-level clouds.
12) I don't see any convincing evidence that an indirect aerosol effect is present. This is speculation on the part of the authors.

Our purpose is to provide some sketches on the possible mechanism that links aerosol production to cloud top height, mediated by water vapor. This is achieved mainly calling on papers in the literature and with some preliminary investigation. This complicated topic deserves more in-depth analysis and we would postpone it for a later publication.

As correctly pointed out by the referee, it is not about an indirect aerosol effect (which has a precise meaning), but a semi-direct aerosol effect (as defined in the IPCC AR4, 2007) instead. Beside the study cited in the manuscript, arguments supporting a decrease in surface insulation of the Indian Ocean have been reviewed, among others, by [Turner 2012].

In Fig. 2, the trend in cloud fraction anomalies is provided together with breakups in summer (JJA) and winter (DJF) months. Fig. 2 suggests a constant decreasing tendency of cloud fraction, which is consistent with values of CF change derived by other instruments ([Stubenrauch et al. 2013, Fig. 3.11, p. 141]) for the same region. Looking at the breakups for JJA and DJF, cloud fraction is not correlated with the seasonal Indian monsoon because its trends over R1 and R2 are commensurate and have equal sign, which wouldn't be otherwise. This finding complies with [Norris, 2001]. In addition, there are indications that high values of water vapor volume mixing ratio (and not number density) are found over desert areas, where moist air is advected upward in presence of biomass burning particles ([Kim et al., 2009]).

The question whether a change in cloud parameters (fraction, optical thickness and top height) translates into a change in columnar water vapor (due to the assumption in the algorithm of ghost column under the cloud) has been recently addressed by [du Piesanie et al., 2013]. They show that water vapor columns derived from SCIAMACHY are influenced neither by changes in cloud fraction nor in cloud optical thickness ([du Piesanie et al., 2013, Fig.4, p.2930]). The dependence of H_2O total column on changes in cloud top height has been explored for cases with CF $\geqslant 0.9$ ([du Piesanie et al., 2013, Fig.5, p.2930]). Even for these very cloudy scenes, a change in CTH of $\approx$450 m over 17 years can't explain the decrease of $\approx$ -0.85 g/cm^{-2} above Arabia for the same time span. This argument rules out algorithmic artifacts due to clouds and points to a real process, that presumably takes place over the northern Indian Ocean.

(Action, Sect. 4.2, l. 342–363) The following paragraphs have been added/rewritten.

"Further arguments supporting a decrease in surface insulation of the Indian Ocean have been reviewed, among others, by [Turner 2012].

In general, changes in column-averaged cloud top height might be explained by changes in cloud cover in different altitude layers. When looking at the trend in annual cloud fraction anomalies (not shown here) a constant decreasing tendency is suggested. This is consistent with long-term changes derived by other instruments [Stubenrauch et al. 2013] for the same region. In the bottom row of Fig. 14 the breakups in summer (JJA) and winter (DJF) months are portrayed. Cloud fraction trends don't exhibit seasonality and are not correlated with the seasonal Indian monsoon because the trends over R1 and R2 are commensurate and have equal sign, which wouldn't be otherwise. This finding also complies with [Norris, 2001].

Moreover, the question whether a change in cloud parameters (fraction, optical thickness and top height) translates into a change in columnar water vapor (due to the assumption in the algorithm of ghost column under the cloud) has been recently addressed by [du Piesanie et al., 2013]. They analysed SCIAMACHY water vapor columns and cloud products, generated with the same algorithms of the datasets used in this work (AMC-DOAS, Noel et al., 2004, Noel et al., 2005; SACURA, Rozanov and Kokhanovsky, 2004). The authors showed that water vapor columns are influenced neither by changes in cloud fraction nor in cloud optical thickness. The dependence of H_2O total column on changes in cloud top height has been explored for cases with CF $\geqslant 0.9$. Even for these very cloudy scenes, a change in CTH of ≈ 450 m over 17 years can't explain the decrease of ≈ -0.85 g/cm^{-2} above the Arabian Peninsula for the same time span. This argument rules out algorithmic artifacts due to the shielding of water vapor by clouds and points to a real process, that presumably takes place over the northern Indian Ocean."

(Action, Fig. 14, page 27) Two new maps have been added to the figure (bottom row), that portray cloud fraction trends at a continental scale, broken up for JJA and DJF.

(Action, Conclusions, l. 400–407) The following paragraph has been rewritten.

"Preliminary analysis of the trends seen over North Africa and Arabian Peninsula suggests a semi-direct aerosol effect on clouds. Natural and anthropogenic soot particles have an impact on cloud top height, modulating insulation and water evaporation over the Indian Ocean. Subsequently, during the winter phase of the Indian Monsoon, transport of moist air masses over the arid Arabian Peninsula exerts a negative feedback and decreases statistical significance of clouds' vertical displacement. In fact, this effect has been already seen over Europe during the economic, industrial and infrastructural adjustments following the fall of the East Bloc (Devasthale et al., 2005). However, this topic deserves in-depth investigation, which we defer to a later publication."

Answer to Anonymous Referee #2

1) The title is misleading, because the presented observations are only capable to determine cloud top height of a sub-sample of clouds: those which are optically thick (cloud optical depth > 5) which correspond probably to about slightly less than half of all clouds. Especially many high-level clouds are missed, because these are mostly semi-transparent. It would have been helpful for the reader to present this fact in the beginning, with a cloud fraction for low-level, mid-level and high-level clouds.

In fact, in the abstract (line 7) is stated upfront that the algorithm is based on the retrieval of cloud properties for "for optically thick clouds". This is because, within the asymptotic theory of radiative transfer, direct sunlight is neglected and the diffusive limit is reached at cloud optical thickness equal 5.

Moreover, the spectrometers used in this work, despite their known limitations, offer some unique features (such as long-term coverage and spectral resolution) and we think that our analysis is still a useful contribution to the task of cloud remote sensing and climate modeling, exactly like every other cloud data set available.

Indeed, as acknowledged by the referee, our data sets are interesting to explore thick clouds, not all clouds. Thus, in this spirit, we hope to provide our results to the scientific community, which will enrich and debate them.

(Action) No change.

The word 'trends' in the title is also misleading, because the authors show that within the uncertainties no linear trends can be found.

[Chandler and Scott, 2011] offer the following definition: "Trend is long-term temporal variation in the statistical properties of a process, where 'long-term' depends on the application". Among all statistical properties which describe time series, for the mean value and its temporal change the word 'trend' has been also proposed in past textbooks ([Chatfield 2003] and [Kendall and Ord 1990]). Therefore, it seems that trend can be used even if significance has not been found. Moreover, the length of this data set (17 years) exceeds the length of other data sets, whose analysis has been already termed trend analysis (e.g., [Marchand 2013]).

(Action) No change.

2) The authors have shown in an earlier paper (Fig. 4, Lelli et al. 2012) that a bias in cloud top height is still optical depth dependent for optical depth larger than 5, especially for ice clouds, when water clouds are situated beneath (which happens quite often according to CALIPSO-CloudSat analyses, especially in the tropics).

These cases are excluded from the analysis, owing to the forward model which assumes single-layer clouds. The text accompanying [Lelli et al., 2012, Fig. 4] reads (Sec. 3.2, p. 1556): "Given that our model assumes single-layered clouds, we would then reject retrievals flagged -3-, above a limit height of 5 km".

The algorithm flags -3- those retrievals exceeding the operational limit of geometrical thickness of 11 km and indicates the heterogeneity of the cloud scene (as demonstrated by [Rozanov et al., 2004]). In addition, it is clearly stated that retrievals flagged 3 (for a CTH $>$ 5 km) are discarded (Sec. 4.1, p. 1561, "Data selection" in [Lelli et al., 2012]).

For an additional quantitative proof that the CTH bias is not correlated with COT error, see the answer to the interactive comment by van Diedenhoven.

(Action) No change.

3) The data used from three different instruments have different foot print sizes, with a quite coarse spatial resolution. Especially the foot print size of GOME (320 km x 40 km) does not seem to be adequate to study low-level clouds, because these may appear at smaller horizontal extent. In this case a decrease in height might be linked to a decrease in horizontal extent of low-level clouds within the foot prints.

Despite the coarse footprint of GOME (as correctly pointed out by the referee), we are persuaded that the instrument has potential for the assessment of CTH long-term changes, given that CTH has been deseasonalized separately for each instrument. Our opinion is based on the following evidences:
(a) the good overlap at the beginning of 2003 with the anomaly time serie of SCIAMACHY, which has a considerably finer footprint;
(b) the GOME capability to detect the ENSO event in 1997-1998, just as good as the GOME-2 (which is again a finer-grained instrument) sensitivity for the ENSO event in 2011.

(Action) No change.

4) When using different instruments for trend analysis, calibration is also important as already indicated by B. van Diedenhoven in his interactive comment.

Please, see the appropriate answer to the short comment by van Diedenhoven.

(Action) No change.

Points 1-4 indicate that it will be very difficult to use these datasets for a linear trend analysis (and why should cloud trends be linear?) in cloud height. Indeed, the authors show that within the uncertainties no linear trends can be found. The maps with trends and their significance (Figs 11-13) are difficult to believe, considering the possible optical depth, vertical structure and foot print size dependent biases listed in points 1-3.

Clearly, we disagree with the conclusions given by the referee. While is not completely true that no trends have been found, because physically consistent patterns can be detected on a regional scale, we think that our retrievals are accurate enough with respect to: optical depth dependency (pt. 2 of this answer; answer to interactive comment by van Diedenhoven); calibration errors (answer to interactive comment by van Diedenhoven); and spatial resolution issues (analysis of anomalies).

In addition, we would like to provide more insight on the suitability of a linear model for the assessment of trends in CTH, despite the fact that it might be out of scope for this work.

We note that the condition of normality is reasonably satisfied even in the region of highest autocorrelation (i.e. Central-East Pacific, Niño 3.4 index region), as can be seen in the left plot of Fig. (3), where an analytical gaussian PDF (red curve) has been fitted to the trend distribution. The right plot of Fig. (3) displays the theoretical residual quantiles (estimated after regression with the linear model and application of the parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ of Eq. 3 of the manuscript) against the sample distribution of global CTH anomalies. Since the majority of points cluster about the straight line, the linear model seems to be a reasonable assumption, given that the autocorrelation functions of Fig. 4 of the manuscript drops almost to 0 after one month.

Therefore, the major process that still can influence the time serie is the quasi-stationary ENSO. [Laken et al., 2012] indirectly came to similar conclusions, while analysing the effect of solar activity on cloud altitudes with MODIS Terra and Aqua measurements (on the decadal time scale). [Norris 2005] comes to the same conclusion as well. The ENSO periodicity is still matter of on-going research (as an example, see [Solomon and Newman 2012]) and it is not the focus of the present paper.

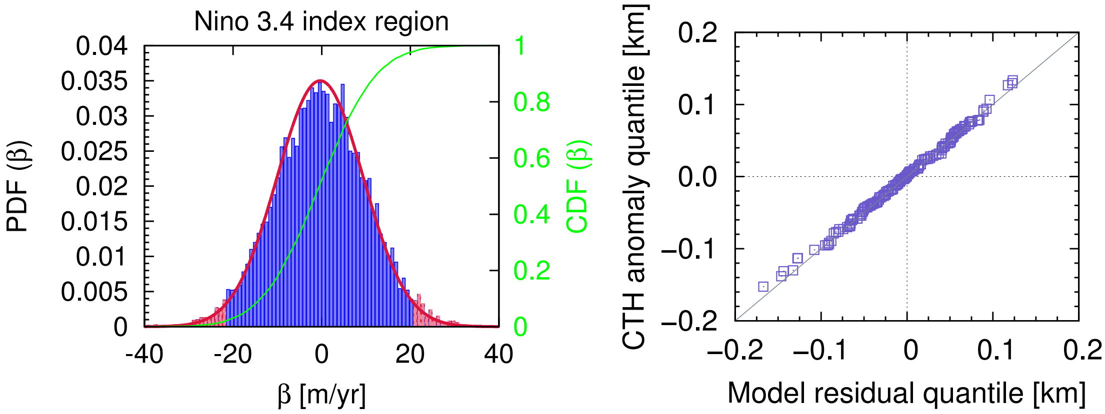

Figure 3: (Left) Trend PDF for the Central-East Pacific region, where residual autocorrelation (≈ 0.12-0.13) has been found. (Right) Normal Q-Q plot for global CTH anomalies.

(Action, Title) The word "Linear" has been added to the title.
(Action, Fig. 5, page 21) The left plot of Fig. 3 has been added to the figure with the trend PDF over the Central East Pacific
(Action, Fig. 6, page 21) The right plot of Fig. 3 has been added as new figure.
(Action, Sect. 3, l. 202–213) The following paragraphs have been added.

"We note that the condition of normality is reasonably satisfied even in the region of highest autocorrelation (i.e., Central East Pacific, 170°W°-120 W°, 5 N°-5 S°), as can be seen in the right plot of Fig. 5. Moreover, Fig. 6 displays the theoretical residual quantiles, estimated after regression with the linear model and application of the parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ of Eq. 3 against the sample distribution of global CTH anomalies. Since the majority of points cluster about the straight line, the linear model seems to be a reasonable assumption, given that the autocorrelation functions of Fig. 4 drops almost to zero after one month. Therefore, the major process that still can influence the time serie is the quasi-stationary ENSO.

Laken et al. (2012) indirectly came to similar conclusions, while analysing the effect of solar activity on cloud altitudes with MODIS Terra and Aqua measurements. Norris (2005) comes to the same conclusion as well. The ENSO periodicity is still matter of on-going research (Solomon and Newman, 2012) and it is not the focus of this work."

Answer to B. van Diedenhoven

I would like to provide a specific comment to the paper "Trends in cloud top height from passive observations in the oxygen A-band" by L. Lelli, A. A. Kokhanovsky, V. V. Rozanov, M. Vountas, and J. P. Burrows (acpd-13-31409-2013). It is not my intent to provide a full review of the paper.

In Figure 1 of their paper, Lelli et al. show time series of cloud top heights retrieved using oxygen A-band measurements of GOME, SCIAMACHY and GOME-2. Offsets between the time series are apparent. The authors hypothesize that this may be due to differences in sensor footprint sizes.

However, I think the presence of calibration errors is a far more likely explanation for these offsets. In our paper van Diedenhoven et al. (2005), we tested the calibration of GOME and SCIAMACHY measurements in the oxygen A-band by retrieving surface pressures in cloud-free conditions and validating them using meteorological data. Over a wide range of surface albedos, a consistent positive bias of 20 hPa in the SCIAMACHY results, compared to those of GOME, was apparent in the data. Furthermore, the GOME results agreed much better with the validation set after accounting for the effects of aerosol on the retrievals. For an average cloud height of 6 km and a atmospheric scale height of 7.4 km, a 20 hPa offset would translate in a ?0.3 km negative bias in the SCIAMACHY cloud top height, which seems very consistent with the offsets in the global results shown in Figure 1.

In van Diedenhoven et al. (2005), we interpreted this offset as resulting from an off-set (rather than scaling) bias in the oxygen A-band measurements and recommended adding 0.86% of the continuum reflectance at 756 nm to the SCIAMACHY reflectance measurements in the oxygen A-band. This correction is apparently consistent with a correction advised by Noël (2004) at the University of Bremen. If these conclusions are accepted it appears that GOME-2 has a greater calibration offset than SCIAMACHY. I would recommend taking such calibration errors into account when interpreting the data.

We acknowledge that in the manuscript no information was given on the version of L1 data used for the generation of the L2 cloud record. For instance, at the time of processing, the ingested SCIAMACHY L1 data were version 7.03, consolidation degree U (year 2010). These L1 data already contain the necessary radiometric key data for correction of calibration offsets pointed by van Diedenhoven et.al, 2005. Please, look at the newest version of the applicable technical note [Noël 2005]) for the actual key data. In its revised form, the manuscript will contain information on the version of L1 data as well as the version of cloud fraction used.

Additionally, we provide Fig. 4. Here the SACURA algorithm's sensitivity to three different calibration errors is portrayed. The input radiance in the O_2 A-band represents a single-layer cloud placed at 5 km (top) and 4 km (bottom altitude) for a optical thickness of 20. COT is set equal to 20 because its global distribution peaks about this value (see [Lelli et al., 2012, Fig. 19, p. 1565]). Solar zenith angle is set equal to 60° with a dark underlying surface.
From top to bottom of Fig.4, the relative error (%) in COT, the absolute error [km] in CTH and CBH are plotted, respectively. For the black curves, the radiance is perturbed only at λ=758 nm. For the red curves, a constant offset has been added in range 758-772 nm, meaning a shift of the whole band. For the blue curves, a wavelength-dependent offset has been added to the whole band. The spectral behavior of the calibration error has been taken from [Noël 2005, Fig. 9, p. 35] and is considered linear from λ=758 nm ($\pm$0.86%) throughout λ=772 nm ($\pm$0.80%). Clearly, this last error parameterization is also the most realistic, as compared to the single-channel per-

turbation in the continuum outside the band.

From Fig. 4, some conclusions (relevant to the scope of this paper) can be drawn:

(1) Spectral-dependent calibration errors have almost no impact in the retrieved CTH. The CTH bias (mid plot, blue curve) is stable about a value of ≈ 250 m, which is the error introduced by the forward model (please, see [Lelli et al., 2012, Fig. 1, p. 1556]). This is a feature of the algorithm, which is based on spectral ratios and on the concurrent fit of CTH and CBH.
(2) In case of calibration errors, CBH is the influenced parameter. The bottom plot clearly shows that CBH becomes noisier.
(3) The COT error doesn't affect the retrieved CTH.

Therefore, we think that the primary role in the shifts among time series of absolute values (Fig. 1 in the manuscript) is played by the different spatial resolution among the instruments and not by radiometric calibration.

(Action, Sect. 2, l. 59–68) The following paragraph has been added.

"At the time of processing, GOME L1b data were in their version 4.00 and SCIAMACHY L1b version 7.03 with consolidation degree U. GOME-2 L1b data for time window 01/2007–12/2009 were in reprocessed stage, while near-real time data have been used for 01/2010–05/2012. Cloud fraction is obtained from broadband Polarization Measuring Devices (PMD) measurements with the Optical Cloud Recognition Algorithm (OCRA) (Loyola and Ruppert, 1998). For GOME, cloud fraction is delivered in bundle with L1b data, while cloud fraction for GOME-2 is taken from the off-line L2 reprocessed dataset made available by DLR in the framework of EUMETSAT's O3M Satellite Application Facility. Cloud fraction for SCIAMACHY has been calculated at University of Bremen with the in-house OCRA implementation."

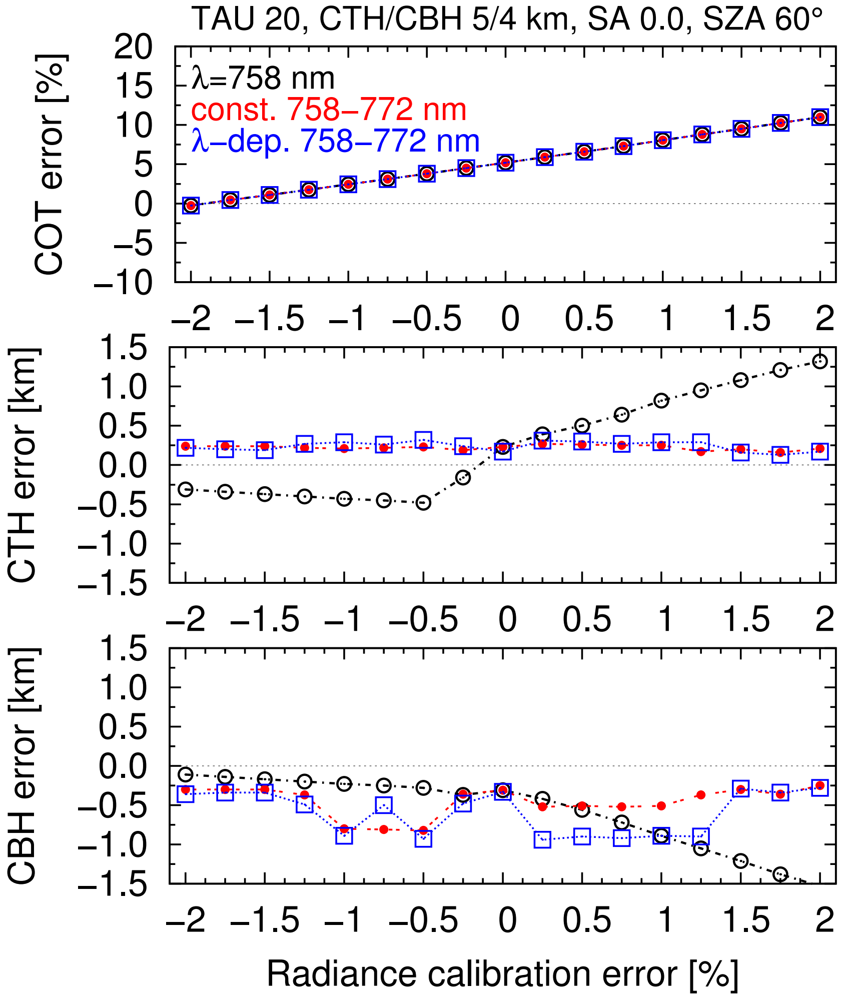

Figure 4: Errors in retrieved cloud optical thickness (%, top plot), cloud top and bottom height (km, mid and bottom plot) for a single-layered cloud, placed at 5-4 km altitude and optically dense 20, above a dark surface. Three different offsets in radiance are applied: single-channel at λ=758 nm (black curves); constant, added along the whole band (red curves); spectral-dependent, as described in [Noël 2005](blue curve).

The following footnotes have been added to **(Tab. 1, page 17)** about instrumental details.

Table 1: Technical specifications of the instruments used in this work. The bottom part of the table describes the Polarization Measuring Devices (PMD).

	GOME	SCIAMACHY	GOME-2
Data availability	1996 – 2011[a,b]	2002 – 2012[c]	2007 – 2022[d]

[a] Global coverage lost in May 2003. [b] Payload has been switched-off since July 2011. [c] Lost contact on April 8, 2012. [d] Foreseen extension of GOME-2 records aboard Metop-B/C.

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 2: 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 5 shows the occurrences of the quality flags (see Table 2 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 3. Table 3 shows also that the number of usable retrievals increases and the relative number of rejections drops as function of instrument. In Figure 6 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.

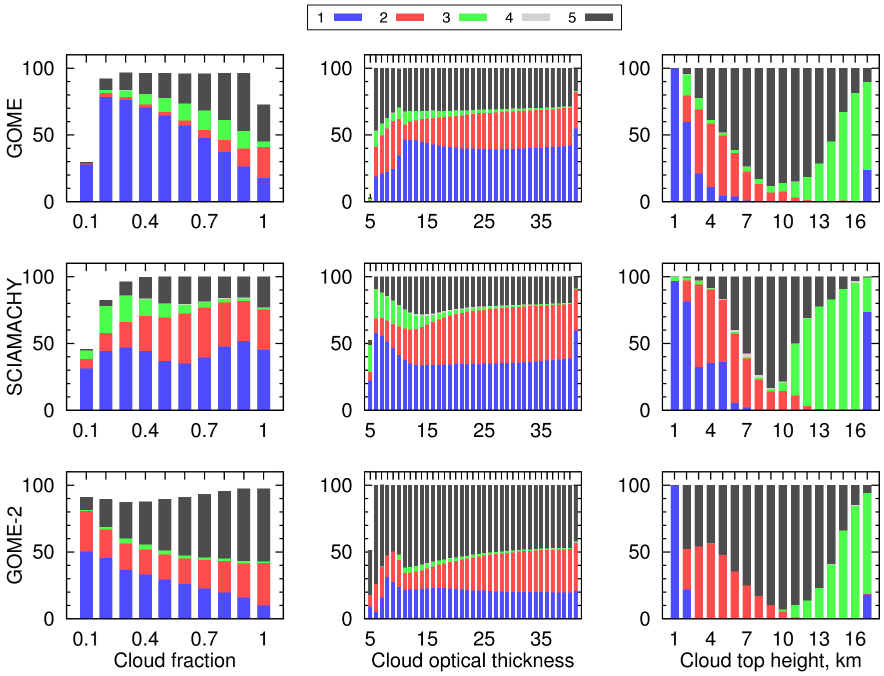

Figure 5: Quality flag statistics of cloud property retrievals used for the analysis, normalized to the total number of retrievals. Note that missing values in the histograms stand for quality flag 0 (“no retrieval”) and aren’t written to the output. The studied sub-sample is populated with black, red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

Table 3: 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

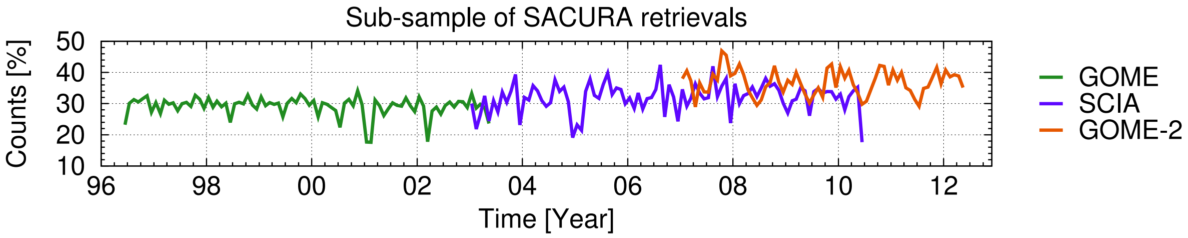

Figure 6: Time series of relative counts in the latitude belt ±60° used for the calculation of the anomaly time series.

References

[Lelli et al., 2012] Lelli, L., Kokhanovsky, A. A., Rozanov, V. V., Vountas, M., Sayer, A. M., and Burrows, J. P.: Seven years of global retrieval of cloud properties using space-borne data of GOME. Atmos. Meas. Tech., 5, 1551-1570, doi:10.5194/amt-5-1551-2012, 2012.

[Mieruch et al., 2008] Mieruch, S., Noël, S., Bovensmann, H., and Burrows, J. P.: Analysis of global water vapour trends from satellite measurements in the visible spectral range, Atmos. Chem. Phys., 8, 491-504, doi:10.5194/acp-8-491-2008, 2008.

[Turner 2012] Turner A.G and Annamalai H.: Climate change and the South Asian summer monsoon, Nature Climate Change 2, 587-595, doi:10.1038/nclimate1495, 2012

[Kokhanovsky et al., 2007] Kokhanovsky, A. A., Mayer, B., Rozanov, V. V., Wapler, K., Burrows, J. P., and Schumann, U.: The influence of broken cloudiness on cloud top height retrievals using nadir observations of backscattered solar radiation in the oxygen A-band, J. Quant. Spectrosc. Rad. Tran., 103, 460–477, doi:10.1016/j.jqsrt.2006.06.003, 2007a.

[Wilks 2011] Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Volume 100, Third Edition (International Geophysics), Academic Press, 2011.

[Stubenrauch et al. 2013] Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H. Getzewich, B., Di Girolamo, L., Guignard, A., Heidinger, A., Maddux, B., Menzel, P., Minnis, P., Pearl, C., Platnick, S., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, Bull. Amer. Meteorol. Soc., pp. 0003Ð0007, doi:10.1175/BAMS-D-12-00117, 2013.

[Norris, 2001] Norris, J. R.: Has northern Indian Ocean cloud cover changed due to increasing anthropogenic aerosol?, Geophys. Res. Lett., 28, 3271Ð3274, doi:10.1029/2001GL013547, 2001.

[Kim et al., 2009] Kim, S.-W., Chazette, P., Dulac, F., Sanak, J., Johnson, B., and Yoon, S.-C.: Vertical structure of aerosols and water vapor over West Africa during the African monsoon dry season, Atmos. Chem. Phys., 9, 8017-8038, doi:10.5194/acp-9-8017-2009, 2009.

[du Piesanie et al., 2013] du Piesanie, A., Piters, A. J. M., Aben, I., Schrijver, H., Wang, P., and Noël, S.: Validation of two independent retrievals of SCIAMACHY water vapour columns using radiosonde data, Atmos. Meas. Tech., 6, 2925-2940, doi:10.5194/amt-6-2925-2013, 2013.

[Marchand 2013] Marchand, R.: Trends in ISCCP, MISR, and MODIS cloud-top-height and optical-depth his- tograms, J. Geophys. Res.-Atmos., 118, 1941Ð1949, doi:10.1002/jgrd.50207, 2013. 31417, 31420

[Chatfield 2003] Chatfield, C.: The Analysis of Time Series - An Introduction, 6th edition. Chapman & Hall/CRC Press, Boca Raton, Florida, 2003

[Chandler and Scott, 2011] Chandler, R. and Scott, M.: Statistical Methods for Trend Detection and Analysis in the Environmental Sciences, Wiley, 2011.

[Kendall and Ord 1990] Kendall, M. and Ord, J.: Time Series, 3rd edition. Edward Arnold, 1990.

[Laken et al., 2012] Laken, B., Palle, E., and Miyahara, H.: A Decade of the Moderate Resolution Imaging Spectroradiometer: Is a Solar-Cloud Link Detectable?, J. Clim., 25, 4430-4440, doi:10.1175/JCLI-D-11-00306.1, 2012.

[Norris 2005] Norris, J.R.: Trends in upper-level cloud cover and surface divergence over the tropical Indo-Pacific Ocean between 1952 and 1997, JGR, 110, D21, D21110, doi=10.1029/2005JD006183, 2005.

[Solomon and Newman 2012] Solomon, A. and Newman, M.: Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record, Nature Climate Change, 2, 9 691–699, doi:10.1038/nclimate1591, 2012.

[Rozanov et al., 2004] Rozanov, V. V., Kokhanovsky, A. A., and Burrows, J. P.: The determination of cloud altitudes using GOME reflectance spectra: multi-layered cloud systems, IEEE T. Geosci. Remote, 42, 1009-1017, doi:10.1109/TGRS.2004.825586, 2004.

[Noël 2005] Noël S.: Determination of Correction Factors for SCIAMACHY Radiances and Irradiances, Technote IFE-SCIA-SN-20050203_IrrRadCorrection (Issue 5.2), `http://www.iup.uni-bremen.de/sciamachy/SCIA_CAL/irrad_corr.pdf`, 2005.