Response to Review #1

We thank the reviewer for the detailed comments which helped to improve the manuscript. Most of the suggestions were gratefully taken up. Please note that numbering for figures and tables is changed because now the Appendix is moved into section 2.2 (Methodology of Method 2) by following the other reviewer's suggestion. In addition, calibration results are now presented for each month following both reviewers' comment. We changed the way to compare three methods in the revised version; i.e. we compared regression results instead of mean biases. This is because many targets used for Method 1 have near-zero reflectances, which cause large uncertainties in estimating relative differences between measured and reference reflectances.

However, most of main conclusions of this paper were not changed. One exception is that we now see Meteosat-8 comparable to Meteosat-9 on a monthly basis.

Overall Recommendation

This paper presents the calibration performance of the 0.6 micron channel of the SEVIRI instrument on-board of Meteosat-8 and Meteosat-9, and the 0.7 micron channel on-board MTSAT. The authors evaluate the calibration accuracy against three calibration methods i.e. i) Ray-matching, ii) cloud properties, and iii) deep convection. The focus of the study is in inter-comparing these three calibration methods. The results show that these methods are capable in finding general biases in calibrations. However, the underestimations of reflectances found in this study differ considerably between the three calibration methods i.e., for METEOSAT-9 the underestimation varies between -3.9 and -6.1%, whereas for MTSAT the underestimation varies between -3.9 and –18.1%.

Finding a reliable calibration procedure is of great importance for climate research, and thus justifies the importance of this paper. The desired accuracy to study long term trends, for example in cloud physical properties, is about 2%. The results presented in this paper give a clue on possible approaches, but they do not meet the desired margin of 2% calibration accuracy. Although the authors used satellite data of different years, they did not evaluate the inter-annual variations in calibration. This would be a very interesting exercise, especially if performed using the different calibration methods proposed.

The paper is generally well written, but the presentation of the results can be improved. The dataset used seems to be rather small, which may explain part of the deviations found. The temporal evolution of the calibration is not presented, but would provide very valuable information to users of METEOSAT and MTSAT data. In addition, it would be useful to evaluate the capability of the calibration methods in detecting trends in calibration. The

paper needs major revisions before it can be published.

For the reliable retrieval of meteorological parameters from the satellite observations, a certain level of accuracy of radiometric calibration is required. Especially, long-term monitoring requires more strict calibration accuracy (~2%) than short-term monitoring (~5%). However, many operational calibration systems, particularly not equipped with onboard calibrator, aim at a 5% calibration accuracy for practical reasons. It is already a challenge to meet the 5% uncertainty level. For example, SEVIRI calibrations aim at an order of 5% accuracy for analysis of meteorological variables (Schmetz et al., 2002), and it has been verified that those accuracy requirements are achieved from vicarious calibration methods using ocean and desert targets (Govaerts and Clerici, 2004).

Please note that the mean biases shown in this study (e.g., high bias of MTSAT-1R measurements at near-zero reflectance, and low bias up to –18.1% for reflectance around 1) do not mean the calibration accuracy of the developed methods. Instead, it suggests the degree of bias of the calibration system employed for each satellite, but at most 5% uncertainty level, that are probably caused by incorrect calibration coefficient and offset. This is based on the fact that Method 1 is a direct comparison against the well-calibrated MODIS measurements. The accuracy of Method 3 has already been examined in Sohn et al. (2009) and demonstrated that the uncertainty level is lower than 5%. Therefore, once Method 2 shows comparable results to Methods 1 and 3, we expect that the combined three calibration method can be reliably used to assess the calibration performance of any visible sensor with the accuracy expected from MODIS measurement.

In order to examine the feasibility of such approach as well as assess the calibration performance of Meteosat-8/9 and MTSAT-1R visible channels, the methods were applied with limited volume of those satellite data. It was demonstrated that Method 1 and 2 are comparable to each other. On the other hand, Method 3 shows a noticeable bias off the regression lines obtained from Methods 1 and 2, and we interpret this discrepancy as a saturation of SEVIRI channel when targets are extremely bright around 100% of reflectance. This interpretation is based on the fact that all three methods are in line with each other when they are applied to MTSAT-1R visible channel. Furthermore the obtained results are consistent with other independent studies about Meteosat-8/9 visible channel calibration using cloud targets (Doelling et al., 2004; Jan Fokke Meirink of KNMI, personnel communication, 2009).

The reviewer suggested to examine time evolution of the calibration using longer data sets. We agree on the importance of evaluating the long-term calculations in particular for the climate monitoring, but it may be beyond the topic we mainly focus on, which is the development of calibration method using cloud targets. Instead it deserves a separate study in

the future. However, cooperating reviewer's concerns, we present results on monthly basis to show some temporal variations of the calibration.

In the abstract,

[Old]

Although the three methods were not in perfect agreement, the results suggest that calibration accuracies were within 5 ~ 10% for the Meteosat-8 0.640-μm channel, 4 ~ 9% for the Meteosat-9 0.640-μm channel, and up to 20% for the MTSAT-1R 0.724-μm channel.

[New]

Although the three methods are not in perfect agreement, the results suggest that calibration coefficients of Meteosat-8/9 0.640-μm channels are underestimated by 6–7%, and larger measurement biases are expected at the high end. On the other hand, calibration coefficients of MTSAT-1R are shown to be underestimated by up to 20% in the bright side, and the space offset count should be corrected.

Results in the revised manuscript are modified and given as follows:

[New]

3. Results

3.1. Meteosat-8/9 SEVIRI 0.640-μm visible channels

The measurements of SEVIRI 0.640-μm channels aboard Meteosat-8 and -9 are compared against MODIS 0.646-μm channel measurements by applying Method 1. MODIS-equivalent SEVIRI 0.640-μm channel reflectances are obtained by applying Eqs. (1)−(4). Regardless of the season, all collocated targets are selected from the area of 40°W–40°E, 20°N–20°S, satisfying smaller SZA (≤ 40°), VZA (≤ 40°) and imposed conditions for the ray-matching. In Fig. 7, comparison is made for each month between measured SEVIRI and MODIS-equivalent SEVIRI reflectances. Their associated regression lines are presented as black solid lines, and corresponding regression statistics and the number of used targets are given in the plots. The number of chosen targets is greater than 1000 for each month and scattered patterns generally cover most of reflectances ranging from 0 to 1. The regression slopes are between 0.894 and 0.936 for Meteosat-8 (four upper panels of Fig. 7) and between 0.925 and 0.939 for Meteosat-9 (four bottom panels of Fig. 7), while intercepts are nearly zero (< 0.01) for both Meteosat-8 and Meteosat-9. No significant difference is noted between Meteosat-8 and Meteosat-9. Obtained regression slopes except for July 2004 are in the range of 0.93–0.94 with near zero intercepts, corresponding to 6–7% of low bias of SEVIRI measurements against MODIS measurements.

The calibration results of Meteosat-8/9 visible channels from Method 1 are quite similar to the results from other previous studies based on the ray-matching technique. Doelling et al.

(2004) showed that MODIS 0.646-μm channel radiances are shown to be 8% larger than Metoesat-8 0.640-μm channel radiances. In addition, Jan Fokke Meirink at KNMI (personal communication, 2009) compares SEVIRI and MODIS visible channels after the atmospheric correction, showing MODIS 0.646-μm channel reflectances 7% (6%) larger than Meteosat-8 (Meteosat-9) 0.640-μm channel reflectances. All these results commonly assert low measurement biases of SEVIRI 0.640-μm channels, implying that the operational calibration of Meteosat visible channels using ocean and desert targets (Govaerts et al., 2004) may underestimate the visible channel reflectance.

Meteosat-8/9 SEVIRI 0.640-μm channel reflectances are simulated using MODIS cloud products as inputs to an RTM, and these serve as references for examining SEVIRI 0.640-μm channel measurements (Method 2). In the simulation, only water cloud targets are used to minimize simulation errors associated with nonspherical cloud particles. In Fig. 8, comparisons are made between simulated and measured SEVIRI 0.640-μm channel reflectances for each month. Because of the applied threshold of COT $\geq$ 5 for selecting cloud-only targets, reflectances smaller than 0.2 are not present in the plots. In general, the selected target numbers are smaller than those used in Method 1 and larger temporal variations in the numbers are evident (e.g., as a worst case only 29 targets are available during April 2007). In Fig. 8, regression lines are given as black solid lines and associated statistics are provided. Regression slopes are between 0.880 and 0.909 for Meteosat-8 (four upper panels of Fig. 8), and between 0.912 and 0.940 for Meteosat-9 (four bottom panels of Fig. 8), while regression intercepts are around 0.02 for both satellites. Regression slopes for Meteosat-8 are generally smaller than those for Meteosat-9, but the differences are not statistically confident because of insufficient cloud targets for certain months (e.g., July 2004, January 2007, and April 2007). In Fig. 8, regression lines from Method 1 are also given as grey solid lines for comparing with Method 2 (black solid lines). Although Method 2 generally produces smaller slopes and larger intercepts than Method 1 for overall periods, black and grey lines are mostly overlaid with each other. Decreased slopes in Method 2 seem to be counterbalanced with increased intercepts. Therefore, it is concluded that Method 2 also provides a similar degree (6–7%) of low measurement biases of Meteosat-8 and Meteosat-9 0.640-μm channels, in comparison to MODIS measurements. Slight differences in regression results between Methods 1 and 2 are likely due to the target reflectances larger than about 0.2 in Method 2, causing larger uncertainties in the regression intercepts. Note that results of Methods 1 and 2 are not displayed separately for Terra and Aqua MODIS. It is because differences between Terra and Aqua are considered to be small compared to the degree of uncertainties of each method, although about 2% differences between Terra and MODIS can be expected (Minnis et al., 2008).

DCC targets are selected using SEVIRI window channel measurements, and the reference

reflectances for those selected DCC targets are produced from simulations with characteristic cloud optical properties (Method 3). We found a bundle of points situated over the simulated reflectance around 1, hindering the linear regression between simulated and measured SEVIRI reflectances. Instead monthly frequency histograms of relative differences [(measured – simulated)/simulated × 100%] are provided in Fig. 9. High peaks in frequency are appeared at similar reflectance values throughout all periods. Monthly means of the relative differences are from –9.1% to –8.6% for Meteosat-8 (left panel of Fig. 9), and from –9.0% to –7.4% for Meteosat-9 (right panel of Fig. 9), implying low measurement biases of SEVIRI 0.640-μm channels. In comparison to Method 1 or Method 2 showing 6–7% of low biases, the degree of biases from Method 3 appears slightly stronger. For the quantitative comparison amongst three methods, we apply regression equations obtained in Method 1 and 2 to predict measurement biases at the MODIS-equivalent reflectance around 1 where DCC targets are located. Measurement biases of –7.8% and –6.5% are predicted from Method 1 for Meteosat-8 and Meteosat-9 DCC targets, respectively. On the other hand, –7.7% and –5.9% of measurement biases are predicted from Method 2. These predicted biases are displayed with vertical lines in Fig. 9. It is certain that Method 3 produces systematically larger biases by 2–3%, in comparison to results from Method 1 or 2.

In Fig. 10, results from Method 3 are directly compared with results from Methods 1 and 2. Monthly regression lines from Methods 1 and 2 (shown in Figs. 7 and 8) are given as grey and black solid lines, respectively, while Method 3 results are given with crosses. Each cross in Method 3 results represents a daily average. As shown in Fig. 9, crosses are below the two regression lines, the discrepancy of the DCC results from Methods 1 and 2 appears to be evident with orders of 2–3%. In Sohn et al. (2009), the accuracy of Method 3 is shown to be within a 5% uncertainty level, and therefore discrepancy of Method 3 from Method 1 or 2 may be attributed to uncertainties in Method 3. However, considering that simulations by the DCC method (Method 3) did not show an apparent bias when applied to the well-calibrated MODIS visible channel (Sohn et al. 2009), the discrepancy of Method 3 may be interpreted as the saturation characteristics of SEVIRI visible channels when targets are highly reflective (Yves Govearts at EUMETSAT, personal communication, 2010). Similar saturation characteristics can also be inferred from the inter-satellite calibration results of Jan Fokke Meirink at KNMI (personal communication, 2009), which shows larger biases of Meteosat-8/9 measurements at the high reflectance end. However, a more detailed explanation appears to be beyond the current research scope and thus deserves a separate examination.

3.2. MTSAT-1R visible channel

MTSAT-1R 0.724-μm channel measurements are compared to MODIS 0.646-μm channel measurements using the ray-matching technique (Method 1). Throughout the seven-month period, all collocated targets are shown to be located in the area of 100°E–180°E, 20°N–20°S.

For the chosen targets, measured MODIS channel reflectances are converted to MODIS-equivalent MTSAT-1R reflectances using Eqs. (5) and (6), and these converted MODIS-equivalent reflectances are compared with measured MTSAT-1R reflectances (Fig. 11). In comparison to the SEVIRI results (Fig. 7), the MTSAT-1R 0.724-μm channel exhibits a more scattered pattern. This is probably due to the scan method of JAMI (David R. Doelling of NASA Langley, personal communication, 2010). Moreover, if MTSAT-1R wobbles from the nominal position (140°E, 0°N), then the viewing geometry at each pixel point is correspondingly changed from the nominal value, causing uncertainties in the ray-matching process. On a monthly basis, linear regression results in slopes between 0.777 and 0.802, and intercepts between 0.015 and 0.037. The slopes are much smaller than 1. Furthermore, intercept off the zero causes the calibration uncertainty depending on the magnitude of reflectance. Bias should be much smaller when reflectance is small, for example in case of reflectance < 0.3. In contrast, underestimate of calibrated reflectance becomes worse when the reflectance becomes larger, resulting in about 20% of underestimate at the high end of MODIS-equivalent reflectance.

MTSAT-1R 0.724-μm channel reflectances are simulated with MODIS cloud products, and the simulation results are used to examine MTSAT-1R measurements (Method 2). Since low cloud occurrences are not frequent enough for the calibration in the analysis domain (100°E–180°E, 40°N–40°S), abundant high cloud targets containing ice cloud particles are included in the Method 2 analysis. In the process, more than 300 calibration targets are collected for each month. Compared to Meteosat-8/9 results (Fig. 8), a larger degree of scatterings is noted between simulated and measured reflectances -- Fig. 12. Again, this may be due to scan problems of MTSAT-1R as well as simulation uncertainties of ice cloud targets related to nonspherical habits. In spite of the large scattering, measured reflectances are linearly correlated with simulated reflectances, showing regression slopes of 0.742–0.799, and intercepts of 0.033–0.059. These results are consistent with those obtained from Method 1, showing a near agreement between two regression lines (i.e. grey line vs. black line shown in each diagram of Fig. 12). Therefore, Method 2 also suggests an underestimate of calibrated reflectance up to 20% at the high end of reflectance and incorrect space count offset of the MTSAT-1R calibration.

These results are consistent with Okuyama (2009) results based on ocean-desert-cloud combined targets in which the regression slope between simulated and measured reflectances was around 0.8. However, Okuyama (2009) showed a near zero intercept offset, significantly different from offset results from both Method 1 and Method 2, probably because of more uncertain parameterization of dark ocean targets.

MTSAT-1R visible channel reflectances simulated over DCC targets using Method 3 are compared with measured reflectances. Monthly frequency distributions of the relative

differences of measured reflectances are plotted against simulation results in Fig. 13 [(measured – simulated)/simulated × 100%]. Although much broader frequency distributions are found compared to SEVIRI visible channels (Fig. 9), their peaks are appeared at similar reflectances over the analysis period, suggesting that results are not much variant over the time. Resultant monthly mean differences ranging from –19.8% to –17.0% are in line with biases predicted from Methods 1 and 2; MTSAT-1R measurement biases at reflectance at 1 were calculated to be –18.3% and –18.7% from Methods 1 and 2, respectively -- find seven month mean biases expressed as vertical grey and black lines in Fig. 13. Differences should be small amongst Methods 1, 2, and 3 because the mean position of high peaks of histograms is located near the vertical grey and black lines.

In Fig. 14, results from Method 3 are compared with those from Methods 1 and 2 using scatter plots. Again results indicate that DCC results are in near agreement with what predicted from two regression results, suggesting the measurement biases of MTSAT-1R visible channel to be between –19.5% and –16.7% around a unit reflectance. The consistency found amongst MTSAT-1R results from Methods 1, 2, and 3 strongly suggests that the 2% of underestimate of reflectance by the Method 3 for the SEVIRI visible channels should not be due to the deficiency of Method 3, but due to the saturation of the sensor capability of detecting the high side of reflectance. Furthermore, it is suggested that the large scattered patterns found in MTSAT-1R results from Methods 1 and 2 (two or three times larger than RMSEs of Meteosat-8/9) are not likely due to the water vapor absorption around 0.724 μm. It is because the DCC results showing the same degree of scattering should not be sensitive to the water vapor absorption, as seen in the comparison between Fig. 10 and Fig. 14.

Modified of summary is following:

[New]

4. Summary

In this paper we examined the performance of operational calibration of Meteosat-8/9 SEVIRI 0.640-μm and MTSAT-1R 0.724-μm visible channels using three calibration methods. The first method is based on the ray-matching technique for inter-satellite calibration. MODIS 0.646-μm channel is used as a reference, and reflectances are compared between MODIS and SEVIRI, and MODIS and MTSAT-1R only over oceanic regions. Regression equations are obtained from radiative transfer simulations to convert measured MODIS reflectances into MODIS-equivalent SEVIRI or MTSAT-1R channel reflectances. The results obtained from the ray-matching technique indicate that SEVIRI calibration coefficient is biased low by 6–7%. On the other hand, MTSAT-1R calibration errors appear to vary with the magnitude of reflectance itself because of the incorrect space offset count; a positive bias near zero reflectance turns into negative bias up to –20% in case of reflectance

around 1.0.

The Meteosat-8/9 and MTSAT-1R channel reflectances are simulated using collocated MODIS cloud products, such as COT, particle effective radius, CTT, and CTP as inputs for the radiative transfer model. In the simulation, the LN-ICA method (Oreopoulos and Davies 1998b) is adopted to describe the subgrid variability because the plane-parallel assumption at each grid box could generate simulation errors by 3D radiative effects. Horizontal radiative interaction appears to be negligible as a result of 0.5°-grid spatial averaging, homogeneity checks [STD($R_{0.6}$) $\leq$ 0.1], and the use of moderate SZAs ($\leq$ 40°). Suggested biases in Meteosat-8/9 and MTSAT-1R visible channel calibration appear to be consistent with results from the ray-matching technique since regression results from two methods are mostly overlapped. This implies that 3D effects are effectively taken into account in the LN-ICA method because the LN-ICA method gives consistent results with results derived from MODIS radiance measurements. However, it is also should be noted that performances of Methods 1 and 2 can be related to each other since cloud products used for Method 2 are from same radiances used for Method 1.

Results from these two methods are compared with those derived from the DCC method. It is suggested that Meteosat-8/9 measurements may not be sensitive enough to discretize the reflectance when targets are highly reflective, suggesting a saturation of measured radiances. In contrast, there is no particular pattern showing the saturation for the MTSAT-1R visible channel as shown in two regression lines going through a bundle of DCC-derived points.

Overall, all three calibration methods show an agreement within 2–3% and suggest that the current Meteosat-8/-9 SEVIRI 0.640-μm channels underestimate reflectances by 6–9%. It is also noted that the current MTSAT-1R visible sensor may be subject to biases, depending on the reflectance ranging from +5% at near 0.1 to −20% at near 1.0. Further study is required to examine why the MTSAT-1R shows a diverse error range depending on the target reflectance.

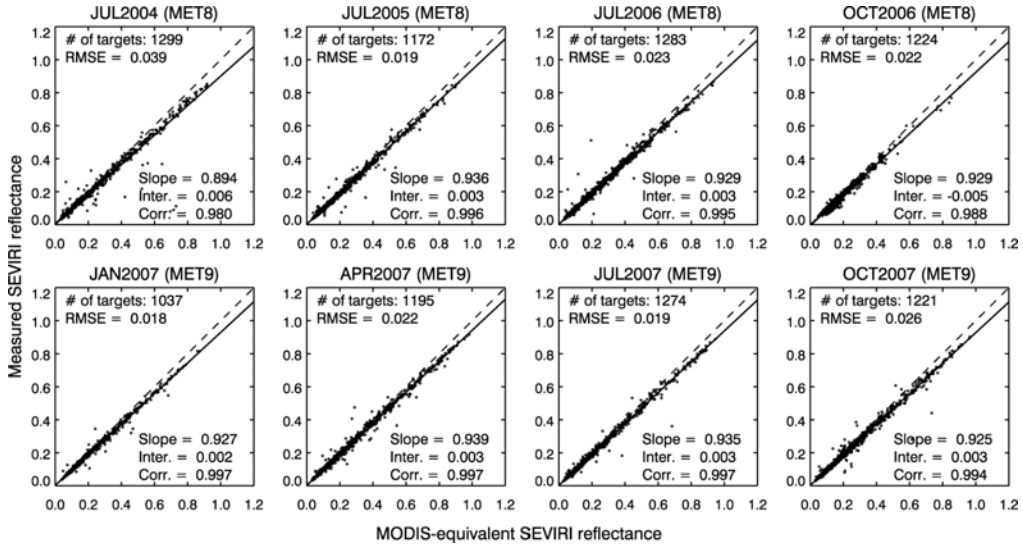

Fig. 7. Scatter plots of MODIS-equivalent SEVIRI vs. measured SEVIRI 0.640-μm channel reflectances of (top) Meteosat-8 and (bottom) Meteosat-9 from Method 1. Regression lines are given as black solid lines along with associated statistics. Dashed lines represent perfect matches.

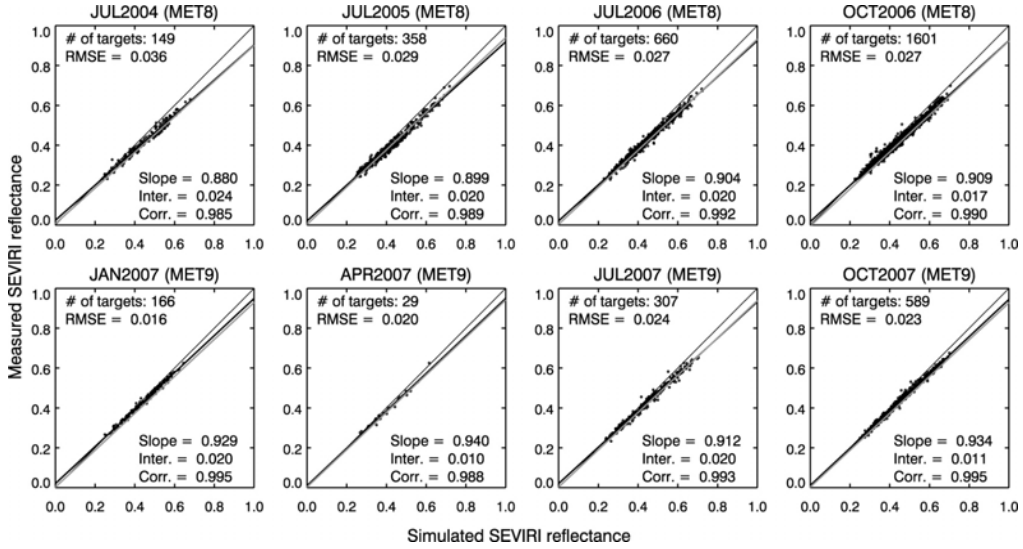

Fig. 8. Scatter plots of simulated vs. measured SEVIRI 0.640-μm channel reflectances of (top) Meteosat-8 and (bottom) Meteosat-9 from Method 2. The simulation is performed for cloud targets using collocated MODIS cloud products. Linear regression results are displayed as black solid lines along with associated statistics. Regression lines from Method 1 are also displayed as grey solid lines.

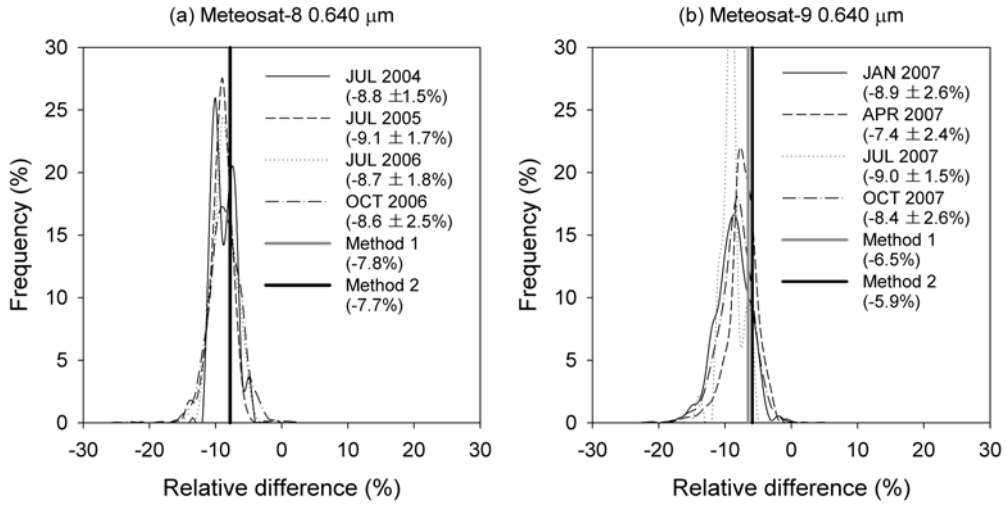

Fig. 9. Monthly frequency histograms of relative differences between measured and simulated reflectances for SEVIRI 0.640-μm channel of (a) Meteosat-8 and (b) Meteosat-9 from Method 3. Mean biases inferred from Methods 1 and 2 are given using grey and black vertical lines, respectively.

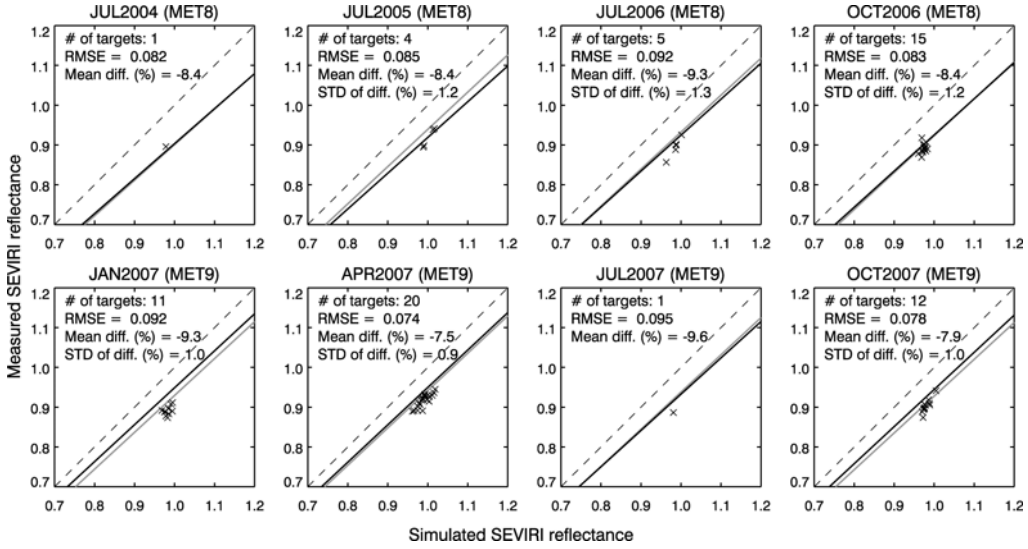

Fig. 10. Comparison of Method 3 (crosses) against Method 1 (grey solid line) and Method 2 (black solid line) for (top) Meteosat-8 and (bottom) Meteosat-9 SEVIRI 0.640-μm channels. For Method 3, the daily average is calculated when the number of selected DCC targets is greater than 10. Dashed lines are perfect matches.

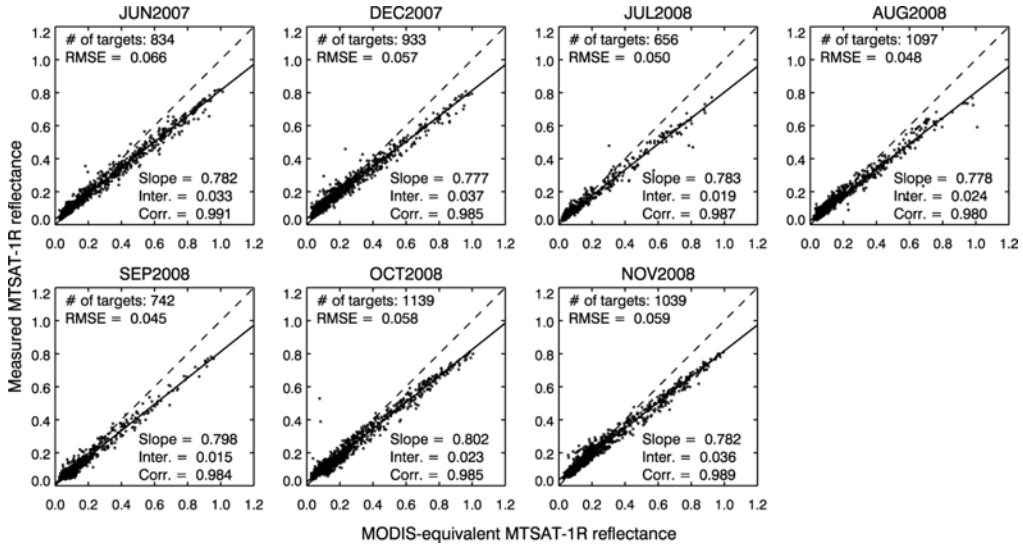

Fig. 11. Same as in Fig. 7 but for the MTSAT-1R 0.724-μm channel (Method 1).

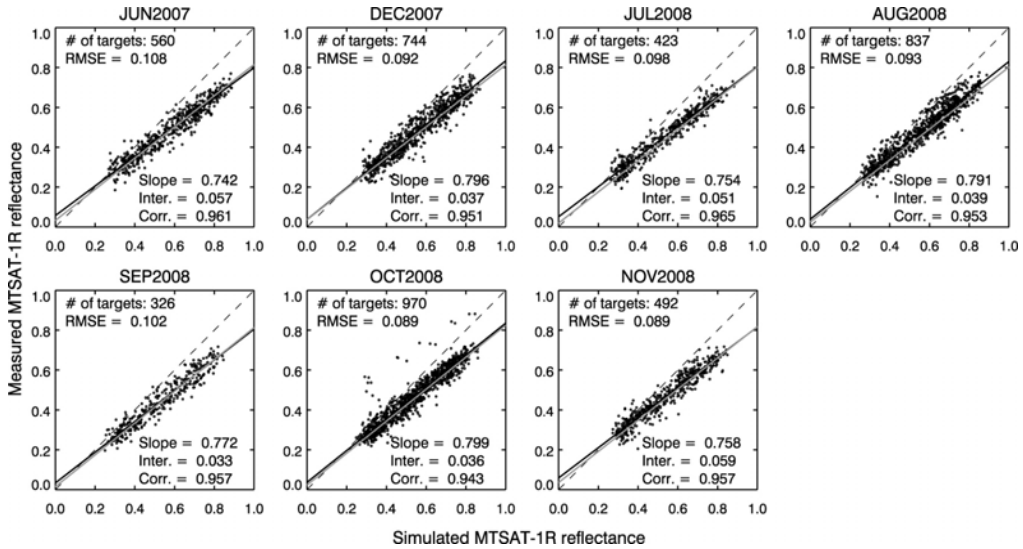

Fig. 12. Same as in Fig. 8 but for the MTSAT-1R 0.724-μm channel (Method 2).

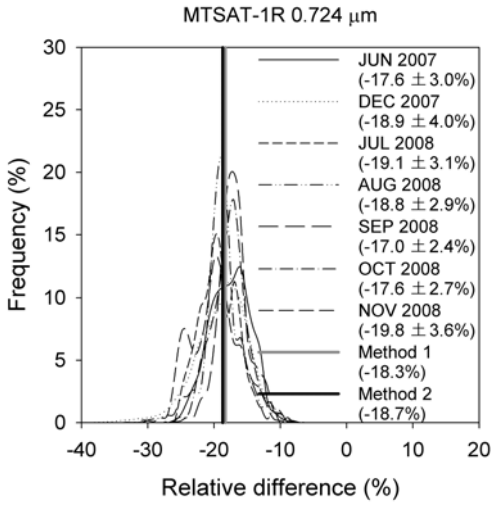

Fig.13. Same as in Fig. 9 but for the MTSAT-1R 0.724-μm channel (Method 3).

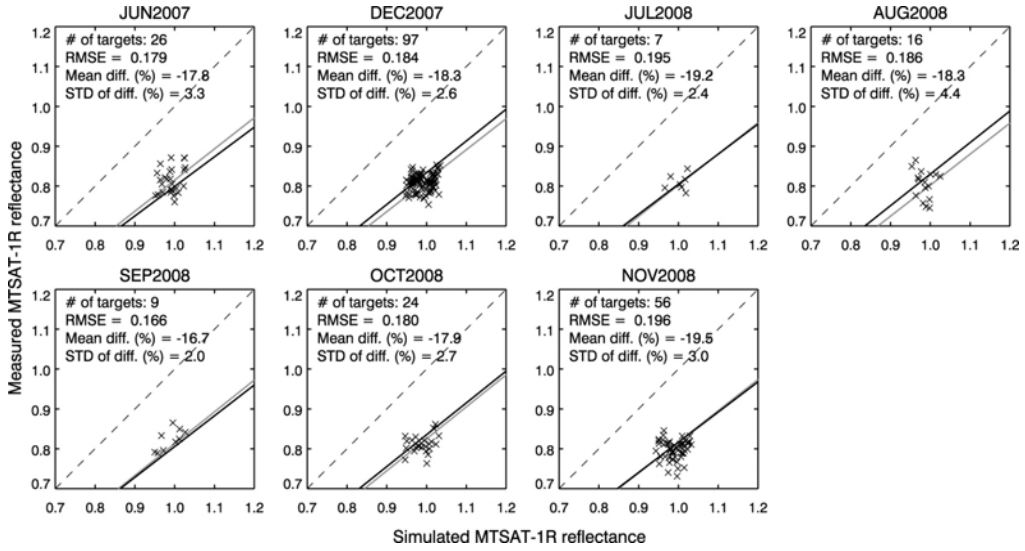

Fig.14. Same as in Fig. 10 but for the MTSAT-1R 0.724-μm channel (Method 3).

Doelling, D., Nguyen, L., and Minnis, P.: Calibration comparisons between SEVIRI, MODIS, and GOES data, in: EUMETSAT Meteorological Satellite Conference, Prague, Czech Republic, pp. 77–83, 2004.

Govaerts, Y. and Clerici, M.: MSG-1/SEVIRI Solar Channels Calibration Commissioning Activity Report, EUMETSAT Technical Note EUM/MSG/TEN/04/0024, EUMETSAT, 2004.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: Radiometric performance of SEVIRI, Bull. Amer. Meteor. Soc., 977–992, 2002.

Sohn, B. J., Ham, S.-H., and Yang, P.: Possibility of the visible-channel calibration using deep convective clouds overshooting the TTL, J. Appl. Meteor. Climatol., 48, 2271–

2283, 2009.

MAJOR CRITICISMS

Point A (Dataset)

• The size dataset seems small. Evaluation period is presented, but it is not clear how much data have been used. This needs to be clarified.

We agree upon the reviewer's concerns of relatively short data period, and we also share your emphasis on the time variation of calibration performance that requires much longer multi-year data sets. However, the main focus of this paper is to develop a method of assessing the visible channel calibration based on the use of cloud targets; here we proposed three different methods. We think several months of data set are sufficient to demonstrate the three different methods to assess the calibration performance. More detailed information on the used data set is now included in the revised version.

[Old]

Four months of SEVIRI data were used for each satellite, i.e., July 2004, July 2005, July 2006, and October 2006 for Meteosat-8, and January 2007, April 2007, July 2007, and October 2007 for Meteosat-9…….. In this study, seven months of data (June 2007, December 2007, July–November 2008) from the 0.724-μm and 11-μm channels were used to examine the visible calibration of MTSAT-1R.

[New]

Four months of SEVIRI data are used for each satellite, i.e., July 2004, July 2005, July 2006, and one month from 15 October to 14 November 2006 (hereafter denoted as October 2006) for Meteosat-8, and January 2007, April 2007, July 2007, and October 2007 for Meteosat-9. For October 2006 case, 32 images per day at 09:00UTC – 17:00UTC are used, whereas for other monthly periods, 8 images per day at 11:00 – 13:00UTC are used. Overall, 1791 and 974 SEVIRI images are used for assessing the calibration status of Meteosat-8 and Meteosat-9, respectively..…….. In this study, seven months of data (June 2007, December 2007, July – November 2008) of 0.724-μm and 11-μm channels are used to examine the visible calibration of MTSAT-1R. Nine images per day at 00:00 – 07:00 UTC and 22:00 – 24:00UTC (total 1926 images) are used.

Point B (Satellite characteristics)

• The position of Meteosat-8 is latitude 0 degrees and longitude -3.4 degrees. Although EUMETSAT provides it's level 1.5 data as if the satellite was positioned at lat 0.0 and lon 0.0, the observed radiances are not modified, and represent the values for lat = 0.0 and lon = -3.4. Did the authors take viewing conditions into account?

Thanks for pointing out. We had found that the position of Meteosat-8 was set as 0° longitude, and thus we corrected data processing with a modified Metoesat-8 position (3.4°W). Changed calibration results are now separately displayed for each month. Modified figures and tables are included in response to overall recommendation.

• The sample size of Meteosat-8 is 3×3 km at nadir, and thus images are provided at a 3×3km resolution at nadir. Indeed, the true sampling size is 4.8×4.8 km. However, taking into account the Spatial Response Function, the majority of the observed radiance comes from a smaller area (see Deneke and Roebeling, 2010, ACPD). Please comment on this.

Now spatial resolution is corrected as 3 km as follows:

[Old]

In this study, the 0.640-μm and 11-μm channels, provided at a 4.8-km spatial resolution, were used to examine the calibration status of the SEVIRI visible channel.

[New]

In this study we use SEVIRI measurements (here 0.640-μm and 11-μm channels) sampled with a spatial resolution of 4.8 km at nadir. However, because of the main signal coming from a small area centered within a 4.8 km pixel and spatial oversampling, the sample size would be smaller about 3 km (Deneke and Roebeling, 2010).

Deneke, H. M. and Roebling, R.: Downscaling of METEOSAT SEVIRI 0.6 and 0.8 micron channel radiances utilizing the high-resolution visible channel, Atmos. Chem. Phys. Discuss., **10,** 10707–10840, 2010.

Point C (Method 3, use of deep convective clouds)

• Method 3. Clouds with CTT < 190 are assumed to have a COT = 200 and re = 20. First, how sensitive are the results to variation in assumed re? Second, I cannot imagine water clouds exist with CTT < 190 k. Third, what is done to ascertain that the anvil is not disturbing the signal? Please comment.

Sensitivity of visible reflectance to the effective radius (r_e) and cloud optical thickness (COT) changes was examined in Sohn et al. (2009). It was shown that up to 3% of reflectance changes are induced by r_e from 10 to 30 μm, while reflectance is varied up to 5% by COT from 100 to 400. Influences of surface reflectance, atmospheric profiles, and aerosol optical depth were also investigated and those are shown to have negligible effects (< 2%) in the DCC simulation.

In the DCC simulation, ice cloud phase (not water phase) is assumed, and thus Baum (2005a, b) scattering model for ice particles is used to specify optical properties of cloud particles. This is because cloud tops of DCCs are generally located near/above the tropopause and upper parts of clouds contain nonspherical ice particles. More detailed description of ice scattering model can be found in Sohn et al. (2009).

In this paper, sensitivity of DCC reflectances to cloud parameters and assumption of cloud thermodynamic phase are now added. Moreover, because another reviewer asked more detailed RTM parameters, now these are also mentioned, i.e.:

[Old]

Once DCC targets were selected, cloud parameters were assumed for the radiative transfer simulation of DCC targets (COT=200 and effective radius=20 μm). Sohn et al. (2009) demonstrated that the simulations of the visible channel reflectance for the DCC targets can be achieved within an uncertainty of 5% using these conditions.

[New]

Once DCC targets are selected, ice cloud phase is assumed since the uppermost part of clouds overshooting the TTL mostly contains nonspherical ice particles. In addition, for the radiative transfer simulation for DCCs, their COT and effective radius are assumed to be 200 and 20 μm, respectively. It is noteworthy to emphasize that MODIS data are not used in Method 3 at all, although *a priori* conditions of COT and effective radius based on MODIS observations are used for the simulation (Sohn et al., 2009). In addition, cloud altitude is assumed to be located between 1 km and 15 km, based on the fact that overshooting clouds are thicker than 10 km (Chung et al., 2008; Luo et al. 2008). Expecting insignificant influence of the atmosphere and surface on the DCC simulation, standard tropical atmospheric profiles and oceanic BRDF model are used, same as in Methods 1 and 2. Note that DCC targets are collected regardless of the land surface types, even though the oceanic BRDF model is used for the calculation of surface reflectance. Sohn et al. (2009) demonstrated that visible channel simulations can be achieved within an uncertainty of 5% using these fixed RTM conditions over DCC targets.

When selecting DCC targets, two types of homogeneity checks and a TB threshold ($TB_{11} \leq$ 190 K) are applied to exclude targets extending to the cloud edge. However, anvil-type cirrus clouds, which are commonly found in the decaying stage of the DCC lifetime, can meet the DCC thresholds because of low cloud top temperatures and smooth cloud morphology by the lateral spreading of DCCs. When applying the DCC method to MODIS visible channel measurements, less than 10% of total selected DCC targets show larger than 5% of calibration errors, probably because of unintended targets such as optically thinner anvil-type cirrus or optically very thick cloud -- Sohn et al. (2009). However, those relatively larger

errors are mostly smoothed out once daily average is taken, showing an accuracy better than 5% on a daily basis. Therefore, we expect that anvil-type cirrus or cloud edges would not produce systematic biases in the calibration results of Meteosat-8/9 and MTSAT-1R visible channels. This is now included in the text as follows:

[Old]

Moreover, two homogeneity conditions were applied to exclude cloud edges. Pixels were selected when the STD of the visible reflectance of the surrounding 9×9 pixels normalized by their mean value was less than 0.03, and the STD of TB_{11} for the same area was less than 1 K.

[New]

Moreover, two types of homogeneity checks are applied to exclude targets extending to the cloud edge. Pixels are selected when the STD of the visible reflectance of the surrounding 9×9 pixels normalized by their mean value is less than 0.03, and the STD of TB_{11} for the same area is less than 1 K. However, anvil-type cirrus clouds, which are commonly found in the decaying stage of the DCC lifetime, can meet the DCC thresholds because of low cloud top temperatures and smooth cloud morphology by the lateral spreading of DCCs. When applying the DCC method to MODIS visible channel measurements, less than 10% of total selected DCC targets show larger than 5% of calibration errors, probably because of unintended targets such as optically thinner anvil-type cirrus or optically very thick cloud -- Sohn et al. (2009). However, those relatively larger errors are mostly smoothed out once daily average is taken, showing an accuracy better than 5% on a daily basis. Therefore, we expect that anvil-type cirrus or cloud edges would not produce systematic biases in the calibration results of Meteosat-8/9 and MTSAT-1R visible channels.

Point D (Sensor degradation)

• *The authors do not present temporal changes in calibration as observed by the different calibration methods. This would be a valuable addition to the results, both from a scientific as from a user's perspective.*

The main objectives of this paper are to develop a calibration algorithm using only cloud targets, and apply the developed algorithm to assess the contemporary SEVIRI and MTSAT-1R visible sensor calibration. We fully agree on the notion that temporal changes in calibration are important, in particular, from the perspective of climate monitoring. We also admit the value of examining the temporal stability of the developed method, although we do not expect significant changes in cloud properties in the foreseeable time period. Such tests with longer data set are, we think, beyond the scope of this paper, and of course deserve a separate study.

MINOR CRITICISMS

Introduction

1. ***Page 4 (line 10):*** *"…minimizing the influence of aerosols in the calculation of radiance at the satellite altitude". This sentence is not clear. Explain this better in the text.*

Now it is corrected as follows:

[Old]

On the other hand, a desert target exhibits small seasonal variations in surface reflection, minimizing the influence of aerosols in the calculation of radiance at the satellite altitude.

[New]

On the other hand, a desert target exhibits strong surface reflection with small seasonal variations, minimizing the relative contribution of aerosols to radiance at the satellite altitude.

2. ***Page 4 (line 15):*** *"…,thus, the intended simulation accuracy may be tolerated with the degree of input accuracies." . This sentence is not clear. Explain this better in the text.*

Now it is corrected as follows:

[Old]

Compared to ocean or desert targets, cloud targets have larger reflectance values; thus, the intended simulation accuracy may be tolerated with the degree of input accuracies.

[New]

Compared to ocean or desert targets, the intended simulation accuracy can be more easily achieved using cloud targets because uncertainties induced by other input parameters are relatively smaller compared to high reflectance of cloud targets.

3. ***Page 5-6:*** *The authors provide information on the observation period for which the re-calibration methods were tested. However, it is not clear how much data were really used in the analysis of each method. The scatter plots suggest that the number of data pairs studied is not large. It should be clarified: how much data are used and if the same overpasses were used for each method.*

The used data amounts are now included in Point A in major criticism. In addition, same overpasses are used for each calibration method. We add following sentence in the last part of data description.

[New]

Note that the same set of Meteosat-8/9 and MTSAT-1R data is equally applied to all three calibration methods.

4. ***Page 5:*** *SEVIRI data are provided at a spatial resolution of 3×3 km (see major comment above).*

Thanks for pointing out. It is now included.

5. ***Page 6:*** *Indicate how the MTSAT and METEOSAT data are synchronized with the MODIS overpass times. Do the authors take into account the time needed to scan a full disk images (about 15 minutes from south to north for METEOSAT).*

Following the reviewer's suggestion, the matching method is now mentioned.

[New]

Matching Meteosat-8/9 with MODIS measurements, the observation time of each MODIS pixel is fixed with one at the center of one granule image. Since the scanning time for one granule is about 5 minutes, the time allocation used in this study can cause at most 2.5 minutes of error. On the other hand, different observation time is applied for each line of Meteosat-8/9 scan image, by considering the full disk image scanned from south to north in about 12 minutes. By the same token, observation time is defined for each line of MTSAt-1R image, but in this case the image is scanned from north to south in 24 minutes.

Methodology

6. ***Page 10 (line 1-10):*** *Replace "former" and "latter" by "first" and "second"*

Thanks for suggestion. Now it is corrected.

7. ***Page 10 (line 25):*** *give the std values that were used to screen on cloud homogeneity in the IR and VIS observations*

Appendix is now moved into section 2.2 (Methodology of Method 2), following other reviewer's suggestion. In the revision, this sentence is removed. However, in the last part of section 2.2, the homogeneity check is mentioned as follows:

[New]

When only spatially homogeneous targets (i.e. CS1 or CS4 types) are selected from applied criteria [$STD(R_{0.646}) \leq 0.1$], the influence of ICA bias for SZA = 40° may be up to 0.005–0.007 with standard deviations of 0.007–0.016 (i.e. ≤2% of absolute reflectances). In conclusion, ICA biases seem to be ignored in Method 2 because of their random distribution and small magnitudes.

8. ***Page 11 (line 1):*** *For the RTM simulations of the DCC targets the authors assume COT=200 and re =20 micron. Here it is not clear if ice or water clouds are assumed. Moreover, it would be good present the sensitivity of the simulated reflectances to assumptions in COT (200 ± 50), effective radius (20 ± 10) and cloud phase (water or ice).*

This is included in Point C in Major criticism. We now mention that more detailed results of sensitivity tests are found in Sohn et al. (2009).

Results

9. ***Fig 1,2,3,4 and 5:*** *Also give number of data pairs and standard deviation of the differences between measured reflectance and simulated or MODIS equivalent.*

Following the reviewer's suggestion, data pairs (# of targets) are provided in scatterplots. In the revised version, we use regression results instead of mean biases (relative differences between measured and reference reflectances) for comparing between Method 1 and Method 2, because of large uncertainties of Method 1 in estimating the mean bias. As a result, mean and standard deviation of differences are no longer provided in results from Methods 1 and 2. In order to give consistent index representing the degree of scattering between measured and reference reflectances for Methods 1, 2, and 3, we decide to provide root mean square error (RMSE) in each diagram. Changed figures are provided in response to the overall recommendation.

10. Include table that summarize the results of the three recalibration methods. Also here bias and standard deviation of the differences.

For example:

Table: Bias and standard deviation of differences for the three re-calibration methods

	Meteosat-8	*Meteosat-9*	*MTSAT*
Method 1	-6.9 % (std_diff??)	-4.1%	+3.9%
Method 2	-5.2%	-3.9%	-13.4%
Method 3	-7.3%	-6.4%	-18.1%

If incorrect space count offset is used in the operational calibration, the regression offset may not be equal to zero. This is evident for MTSAT-1R channel because regression intercept is considerably large (>0.02) and regression slope is around 0.8. As a result, the measurement bias of MTSAT-1R channel is shown to be positive in near-zero reflectance, whereas the bias is up to –20% at high end of the reflectance. Considering that typical reflectance ranges of used targets for Methods 1, 2, and 3 are different, i.e. 0–1 for Method 1, 0.2–0.9 for Method 2, and around 1 for Method 3, direct comparison of inferred measurement biases between

three methods may not be adequate. Therefore, instead of providing above table, we compare three methods in the scatter diagrams in Fig. 10 and 14 to provide features of all three methods.

11. ***Page 13 (line 11):*** *"Note that the results for ray-matching ... around 2~3%" In this sentence the authors argue why a larger bias is expected from the DCC method. However, table 1 shows that the bias of the Meteosat reflectances shows hardly any increase with increasing reflectance (-7.8 to -7.3 (meteosat-8) and -4.0 to –6.4 (meteosat-9). So this table does not explain the higher bias at reflectances close to 1. In addition, it is remarkable that the biases in this table are always larger than the mean biases presented in the scatterplots (-6.9 and –4.0).*

Our view is same as reviewer's. Even if measurement bias is defined with respect to various reflectances as in Table 1, it cannot fully explain higher bias shown in Method 3, compared to Method 1 or 2. Therefore, we interpreted the discrepancy as the saturation characteristics of SEVIRI visible channels. This is discussed in the result, i.e.:

[New]

As shown in Fig. 9, crosses are below the two regression lines, the discrepancy of the DCC results from Methods 1 and 2 appears to be evident with orders of 2–3%. In Sohn et al. (2009), the accuracy of Method 3 is shown to be within a 5% uncertainty level, and therefore discrepancy of Method 3 from Method 1 or 2 may be attributed to uncertainties in Method 3. However, considering that simulations by the DCC method (Method 3) did not show an apparent bias when applied to the well-calibrated MODIS visible channel (Sohn et al. 2009), the discrepancy of Method 3 may be interpreted as the saturation characteristics of SEVIRI visible channels when targets are highly reflective (Yves Govearts at EUMETSAT, personal communication, 2010). Similar saturation characteristics can also be inferred from the inter-satellite calibration results of Jan Fokke Meirink at KNMI (personal communication, 2009), which shows larger biases of Meteosat-8/9 measurements at the high reflectance end. However, a more detailed explanation appears to be beyond the current research scope and thus deserves a separate examination.

As mentioned before, calibration targets with near-zero reflectance exhibit large variations in terms of relative difference in spite of small changes in terms of absolute reflectance. Therefore, it is concluded that targets with near-zero reflectances generate large uncertainties in estimating relative differences. Those large uncertainties likely force mean differences out of ranges that are estimated from regression equation (Table 1). That is why we decided to use regression results for comparing Methods 1 and 2 in the revised version, instead of mean

biases.

Summary

12. Discuss how the findings of this study are related to findings of earlier studies and explain reasons for the observed differences between the methods.

Please note that measurement biases inferred from three methods can be different from each other if incorrect space offset is used in operational calibration, as discussed in point #10. Now obtained results of this study are compared with findings of earlier studies.

In section 3.1,

[New]

The calibration results of Meteosat-8/9 visible channels from Method 1 are quite similar to the results from other previous studies based on the ray-matching technique. Doelling et al. (2004) showed that MODIS 0.646-μm channel radiances are shown to be 8% larger than Metoesat-8 0.640-μm channel radiances. In addition, Jan Fokke Meirink at KNMI (personal communication, 2009) compares SEVIRI and MODIS visible channels after the atmospheric correction, showing MODIS 0.646-μm channel reflectances 7% (6%) larger than Meteosat-8 (Meteosat-9) 0.640-μm channel reflectances. All these results commonly assert low measurement biases of SEVIRI 0.640-μm channels, implying that the operational calibration of Meteosat visible channels using ocean and desert targets (Govaerts et al., 2004) may underestimate the visible channel reflectance.

In section 3.2,

[New]

These results are consistent with Okuyama (2009) results based on ocean-desert-cloud combined targets in which the regression slope between simulated and measured reflectances was around 0.8. However, Okuyama (2009) showed a near zero intercept offset, significantly different from offset results from both Method 1 and Method 2, probably because of more uncertain parameterization of dark ocean targets.

Okuyama, A.: Vicarious calibration of visible channel activity at the JMA, GSICS GRWG-GDWG joint meeting, Tokyo, Japan, January 28-30, 2009.

13. The results of ray-matching in table 1 contradict with the mean biases given in the scatterplots. As mentioned above, this should be explained.

This is included in #11.

14. Page 16 (line 25): "Overall, the three calibration methods showed agreement within 2 to 3% …" This is the case for METEOSAT but not for MTSAT. That should be clarified better.

Three methods are compared in the scatterplots, and fairly consistent results are found for MTSAT-1R visible channel (Fig. 14). On the other hand, for case of Meteosat-8/9 visible channels, Method 3 shows larger biases than what can be expected from Method 1 or 2 (Fig. 10), probably due to the radiance saturation of SEVIRI visible channels at the high end. Nevertheless, the degree of deviation of Method 3 from Methods 1 and 2 for SEVIRI visible channels is within 2–3%.

Annex

15. The difference between the calculation of Rsim(<tau>) and <Robs (tau)> is difficult to understand.

Reflecting the reviewer's opinion, the equation is corrected as follows.

[Old]

$$\Delta R_{PPH} = R_{PPH} - R_{ICA} = R(\langle\tau\rangle) - \langle R(\tau)\rangle \qquad \text{(A1)}$$

where ΔR_{PPH} is the PPH bias; R_{PPH} is the reflectance at a grid from the PPH assumption; R_{ICA} is the reflectance at a grid from the independent column approximation (ICA) considering the subgrid variations; τ is COT; and the operator < > represents a grid average.

[New]

$$\Delta R_{PPH} = R_{PPH} - R_{ICA} = R\left(\frac{1}{N}\sum_{i=1}^{N}\tau_i\right) - \frac{1}{N}\sum_{i=1}^{N}R(\tau_i) = R(\langle\tau\rangle) - \langle R(\tau)\rangle \qquad (7)$$

where ΔR_{PPH} is the PPH bias; R_{PPH} is the reflectance at a grid box from the PPH assumption; R_{ICA} is the reflectance at a grid box from the independent column approximation (ICA) after considering the subgrid variations; τ_i is COT at a subgrid point; N is the number of subgrid points in a grid box; and the operator < > represents a grid average.

16. The authors present in table A1 two methods: the PPH method and the LN-ICA method. The PPH method is mentioned first in the section that presents with the LNICA method. To improve the readability it would be good to also introduce the PPH method also clearly in the previous paragraphs.

Following the reviewer's suggestion in point #15, we change the Eq. (A1) [now Eq. (7)] to clarify how grid reflectances are obtained from PPH and ICA methods respectively. Please note that LN-ICA method is adopted later as an approximate method to calculate Meteosat-8/9 and MTSAT-1R visible channel reflectances. This is because ideal ICA calculation is practically inefficient due to the computational burden. Therefore, it is fine if PPH and ICA methods are introduced first and approximate LN-ICA method later.

17. The authors find no systematic effect due to 3-D cloud structures, This contradicts with the findings of Loeb, Varnai, and Winker, 1998 (JAS), who found an systematic increase of reflectance for bumpy cloud field in backward scattering direction. Please check this.

In this study, 3-D effects are divided into two parts, i.e. the effect associated with horizontal variations and the effect associated with horizontal interactions. Those two effects are examined in terms of PPH and ICA biases. Since PPH bias becomes significantly positive for the 0.5°-grid simulation, LN-ICA method is adopted to resolve sub-grid variations. On the other hand, ICA bias is shown to be randomly distributed, showing both positive (e.g., cloud shadow in forward direction) and negative (e.g., cloud illumination in backward direction) signs, and thus we simply take spatial or temporal average to minimize the ICA bias.

In Loeb et al. (1998), 1D simulation bias is defined as the difference between plane-parallel simulation and Monte Carlo simulation results. Therefore, the simulation bias in Loeb et al. (1998) appear to be the summation of PPH bias and ICA bias of our study. We expect that systematic biases shown Loeb et al. (1998) are partly contributed by PPH bias. Since we resolve sub-grid variability, those systematic biases associated with PPH assumption would be not included in our results. On the other hand, dependence of simulation biases on the viewing geometry is linked to characteristics of ICA bias shown in our study. We also found that ICA bias can be significantly increased as a result of cloud shadow and illumination for certain geometries, but it can be mostly smoothed out from spatial/temporal averaging. In Loeb et al. (1998), special geometry (forward or backward direction) is only focused with a relatively small scale (4-km resolution). Therefore, we expect that the simulation bias shown in Loeb et al. (1998) can be minimized once temporal or spatial average is performed with various viewing and solar geometries, as in our study.

Loeb, N. G., Varnai, T., and Winker, D. M.: Influence of subpixel-scale cloud-top structure on reflectances from overcast stratiform cloud layers, J. Atmos. Sci., 55, 2960–2973, 1998.

Grammatical slips

In general the paper is clearly written. Still some grammatical slips and spelling errors remain. Below some of these are listed.

Thanks for careful reading. Now all points are reflected.

- ***Page 8 (line 19)****: "influence" should be "influences"*
- ***Page 13 (line 3)****: "smaller those observed" should be "smaller than those observed"*