ELSEVIER

Atmospheric Research 72 (2004) 17–37

ATMOSPHERIC RESEARCH

www.elsevier.com/locate/atmos

Validation and homogenization of cloud optical depth and cloud fraction retrievals for GERB/SEVIRI scene identification using Meteosat-7 data

A. Ipe *[,1], C. Bertrand, N. Clerbaux, S. Dewitte, L. Gonzalez

Royal Meteorological Institute of Belgium, Department of Observations, Section Remote Sensing from Space, Avenue Circulaire 3, B-1180 Brussels, Belgium

Received 4 August 2003; received in revised form 30 January 2004; accepted 31 March 2004

Abstract

The Geostationary Earth Radiation Budget (GERB) instrument was launched during the 2002 summer together with the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) on board of the Meteosat Second Generation (MSG) satellite. This broadband radiometer will aim to deliver near real-time estimates of the top of the atmosphere (TOA) radiative fluxes at high temporal resolution thanks to the geostationary orbit. To infer these fluxes, a radiance-to-flux conversion needs to be performed on measured radiances. Since we plan to carry out such a conversion by using the angular dependency models (ADMs) developed from the Clouds and the Earth's Radiant Energy System (CERES) experiment, the GERB ground segment will have to rely on some scene identification on SEVIRI data which mimic as close as possible the one from CERES in order to select the proper ADM. In this paper, we briefly present the method we used to retrieve cloud optical depth and cloud fraction on footprints made of several imager pixels. We then compare the retrieval of both features on the same targets using nearly time-simultaneous Meteosat-7 imager and CERES Single Satellite Footprint data. The targets are defined as CERES radiometer footprints. We investigate the possible discrepancies between the two datasets according to surface type and cloud phase and, if they exist, suggest some strategies to homogenize GERB retrievals based on CERES ones.

Keywords: Scene identification; Cloud optical depth; Cloud fraction; Visible imager

* Corresponding author. Tel.: +32-2-3730631; fax: +32-2-3746788.
E-mail address: Alessandro.Ipe@oma.be (A. Ipe).

[1] Supported by the Belgian Science Policy Office (DWTC/SSTC) through PRODEX-6 contract No. 15162/01/NL/SFe(IC).

0169-8095/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.atmosres.2004.03.010

1. Introduction

To derive the most accurate top of the atmosphere (TOA) solar fluxes from the measured Geostationary Earth Radiation Budget (GERB) (Harries and Crommelynck, 1999; Dewitte et al., 2000) broadband radiances, the Royal Meteorological Institute of Belgium (RMIB) is planning to use the new angular dependency models (ADMs) from the Clouds and the Earth's Radiant Energy System (CERES) experiment (Loeb et al., 2003). They are due to replace and outperform the accuracy of the previous Earth Radiation Budget Experiment (ERBE) models (Suttles et al., 1988). To reach the highest level of confidence in the derived fluxes from the radiance-to-flux angular conversion, it is needed to select and use the corresponding ADM that would be applied by CERES to the measured scene types of each GERB footprint's pixel. This requires the performance of a scene identification (SI) which gives similar results (i.e. features values) than those used to stratify the CERES ADMs. More precisely, it is crucial to remove any systematic bias between these two SI schemes.

As the GERB experiment aims to deliver products on a near realtime basis (i.e. within 4 hours after the acquisition time), the overall RMIB GERB Processing (RGP) should not take more than the time period between two GERB/SEVIRI series of images (i.e. 15 minutes). In order to cope with this major time constraint, such a SI scheme for GERB is committed to remain simple and only extract the smallest needed set of features. It turns out that the minimal set required to select the more relevant ADM is defined by the cloud fraction, the cloud phase, the cloud optical depth and the surface type for each GERB pixel, or more specifically the mean of the feature values of all the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) pixels (resolution of 3 km at nadir) within this GERB footprint's pixel. The use of CERES ADMs in connection with measured footprints of similar spatial resolution than those used to build these ADMs (i.e. about 10 km at nadir, Wielicki et al., 1996) certainly avoids any new error source in the radiance-to-flux conversion which could arise from distinct spatial extends of GERB and CERES footprints. Hence, in the following, the term GERB pixel or footprint will not refer to the native spatial resolution of the instrument (50 km at nadir), but to a pixel's size of about 10 km. To reach such a resolution the GERB broadband measured radiances are being interpolated by some data fusion algorithm with the help of the SEVIRI multispectral imager on board of the Meteosat Second Generation (MSG) satellite (Schmetz et al., 2002). Briefly, the high spectral accuracy of GERB is combined with the high spatial resolution of SEVIRI. This is achieved by correcting the broadband radiance estimations from the 11 SEVIRI spectral channels with GERB measurements and by interpolating these correction coefficients to footprints of 3×3 SEVIRI pixels (Gonzalez et al., 2000).

In this paper, we shall focus on the determination of the cloud optical depth and cloud fraction. We will present a fast cloud optical depth retrieval algorithm which is not based on an iterative scheme as it is the case for the CERES experiment. Instead, it only uses visible narrowband radiances from an imager to estimate the cloud optical depth. Moreover, it does not require to priorly flag cloudy pixels by some cloud detection scheme (generally based on threshold tests applied to imager narrowband measurements). However, look-up tables computed using a radiative transfer (RT) model are still needed in order to link the measured quantities to cloud optical depth, but an innovative way to

parameterize them will be given. The cloud fraction estimation is based on cloudy imager pixels detection within some defined footprint (typically of the size of GERB pixel); the cloudy test being simply a threshold on their associated cloud optical depth. Then, we will illustrate both retrieval schemes by applying them to Meteosat-7 (MS-7) visible images. We will show comparisons of the RGP retrieved cloud optical depth and cloud fraction from Meteosat-7 simulated footprints and CERES associated features' values. We will then suggest some correction schemes on GERB cloud optical depth and provide the optimal threshold for cloudy/non-cloudy pixels which mimic as close as possible the CERES cloud optical depth and cloud fraction estimates. Finally, we will propose some strategies for more in-depth future comparisons of CERES and SEVIRI cloud properties retrievals.

2. Cloud properties retrieval algorithms

2.1. Cloud optical depth

Cloud optical depth retrieval scheme is usually part of a larger cloud properties inference algorithm which is based on multi-wavelength radiance measurements. Such algorithms can generally be divided into two classes: (a) those using RT calculations combined with a threshold test to detect cloudy pixels and (b) those relying on some clustering techniques in the multidimensional radiance space constituted by several bands. However, due to problems inherent to clustering techniques, the majority of cloud remote sensing studies is based on comparisons with RT computations. Nevertheless, clustering schemes are specifically aimed to deal with only partially cloudy pixels, i.e. low spatial resolution measurements, as shown in Arking and Childs (1985).

Threshold techniques treat each pixel independently and assume that they can be considered as either completely cloudy or clear. According to the SEVIRI spatial resolution of 3 km (at nadir), this independent pixel approximation (IPA) can be considered as valid in our case. However, it is worth pointing out that correction schemes exist in order to cope with 3-dimensional cloud structures and their associated horizontal transport of radiation which cannot be anymore neglected at higher spatial resolutions, as it is illustrated in Wyser et al. (2002). IPA algorithms rely on the fact that the reflection function of clouds at a non-absorbing narrowband in the visible part of the electromagnetic spectrum is primarily dependent on the cloud optical depth. These schemes can further be divided into two groups: (1) those making use of asymptotic relations for the reflection and transmission functions of optically thick layers ($\tau \gtrsim 9$) as shown in King (1987) and (2) those using a large set of precomputed values of the reflection function according to various scene viewing geometries and cloud optical depths, τ, as for example in Nakajima and King (1990).

From the previous considerations, it turns out that the simplest algorithm which can still allow the broadest range of applicability for cloud optical depth retrievals is based on simulated radiances from RT models properly convoluted with the visible narrowband filters of the imager under consideration. For simplicity we have simulated the outgoing radiances $L(\theta_0, \theta, \varphi)$ [W m^{-2} sr^{-1}] for a small set of ideal scene types, i.e. pure ground

surfaces with spectrally uniform Lambertian albedos and two cloud conditions (water or ice) with fixed parameters (altitude and particle size). The scene observing conditions are defined relatively to the local normal to the ground surface, where θ_0 is the Sun's zenith angle, θ, is the satellite's zenith angle and, φ, is the relative azimuth angle defined as the angle between the principal plane (made by the Sun, the surface and its local normal) and the satellite-surface-normal plane ($\varphi = 0$ corresponds to forward scattering). We selected for this work the STREAMER (Key and Schweiger, 1998) RT code. The definitions of the scenes simulated by the RT code, as well as its input parameters, are given in Table 1. We used for planetary boundary layer standard atmospheric profile the mid-latitude summer properties and we set the aerosol optical model to maritime for ocean and rural for land, both with background tropospheric and stratospheric vertical profiles and finally we used a 50 km visibility. Gaseous absorption and *Rayleigh* scattering were also included. The cloud optical depths, τ, given at 0.65 μm are chosen according to a pseudo-logarithmic scale. The optically opaque cloudy condition, $\tau = 128$, we used is identical to the one chosen in the CERES experiment, as well as the particle shape for ice clouds. Note that we are only considering single uniform cloud layers in these computations. Moreover, the reflectance model of the ground surfaces used is Lambertian as it is also the case for a majority of algorithms found in the literature (Arking and Childs, 1985; King, 1987; Nakajima and King, 1990; Rossow and Garder, 1993; Nakajima and Nakajima, 1995) and in particular for CERES.

The common approach found in the literature is, first, to apply a cloud detection scheme using threshold tests on multispectral measurements as, for example, the APOLLO algorithm (Saunders, 1986). Then, for so flagged cloudy pixels, RT computed radiances are used directly as look-up tables to extract the cloud optical depth according to the scene viewing geometry and the measured visible narrowband values (see for example Rossow and Garder, 1993). Instead, we proceed differently. As shown in Nakajima and Nakajima (1995), there is an empirical law between the visible radiance (thus reflectance) and the cloud optical depth which is nearly insensitive to the cloud particle size r_e for visible

Table 1
Scene parameters used as inputs for the STREAMER code simulations where α is the broadband surface albedo, z is the bottom cloud altitude, h is the cloud geometrical thickness, r_e the cloud particle mean effective radius and, τ the cloud optical depth at 0.65 μm

α	$0 \rightarrow 0.8$ by 0.05 steps	
Aerosols	maritime	rural
	background tropos. and stratos. prof.	
Std prof.	mid-latitude summer	
θ_0 [°]	$0 \rightarrow 90$ by 5° steps	
θ [°]	$0 \rightarrow 90$ by 5° steps	
φ [°]	$0 \rightarrow 180$ by 10° steps	
Cloud phase	water	ice
z [km]	3	9
h [km]	3	1
Cloud particle Shape	spheric	hexagonal
r_e [μm]	8	70
τ	$0, \{10^{-2}, 10^{-1}, 1, 10\} \times \{1, 2, 4, 7\}, 100, 128$	

wavelengths. By introducing the mean cloud amount (or cloud coverage index), C, defined as (Cano et al., 1986)

$$C(\theta_0, \theta, \varphi, \alpha, \text{phase}, \tau) = \frac{L(\theta_0, \theta, \varphi, \alpha, \text{phase}, \tau) - L(\theta_0, \theta, \varphi, \alpha, \tau = 0)}{L(\theta_0, \theta, \varphi, \text{phase}, \tau = 128) - L(\theta_0, \theta, \varphi, \alpha, \tau = 0)},$$

which can be equivalently expressed in terms of bidirectional reflectance distribution function (BRDF) ρ as

$$C(\theta_0, \theta, \varphi, \alpha, \text{phase}, \tau) = \frac{\rho(\theta_0, \theta, \varphi, \alpha, \text{phase}, \tau) - \rho(\theta_0, \theta, \varphi, \alpha, \tau = 0)}{\rho(\theta_0, \theta, \varphi, \text{phase}, \tau = 128) - \rho(\theta_0, \theta, \varphi, \alpha, \tau = 0)}, \tag{1}$$

where $\tau = 0$, represents clear-sky conditions above the ground surface, $\tau = 128$, denotes the opposite boundary limit associated with optically opaque cloudy conditions leading to simulated radiance fields insensitive to surface albedo. This law is rescaled with C values ranging from 0 to 1. This empirical law is easily built from our RT simulated data. The most noticeable fact when plotting C according to the cloud optical depth is its similar shape for all scene viewing geometries (see Fig. 1). Note that the curve is not saturated at $C = 1$ when $\tau = 128$. This relies on the fact that the reflection function of the optically opaque clouds generated by the RT code is still increasing with the cloud optical depth. Nevertheless, it can be shown that saturation is reached for $\tau \approx 400$. However, to remain in accordance with the CERES retrievals, we have set $\tau = 128$ as the highest possible cloud optical depth value which can be given by our algorithm.

Due to this characteristic shape we parameterize this relation by using a modified sigmoïd function of the logarithm of the optical depth,[2] i.e. $C = f(\log \tau)$, where f is defined as

$$f(\log\tau) = \frac{A}{B + 10^{-(\log \tau - \log \tau_0)/\chi}},$$

and relates on four parameters (A, B, $\tau_0 > 0$ and $\chi \neq 0$). Note that, in the latter expression, the dependency of C and of the four parameters with the scene viewing geometry (θ_0,θ, φ), surface albedo α and cloud phase is implicitly assumed. Some basic algebra leads us to the compact form

$$C = \frac{A}{B + \left(\frac{\tau_0}{\tau}\right)^{1/\chi}} \tag{2}$$

These four parameters can then be adjusted in order to get the closest match between the empirical curve and our modified sigmoïd through a least square fit on the RT data. Practically, the Powell multidimensional fitting method (Press et al., 1992) is used. Finally, we end up with a set of four parameters for each scene viewing geometry, surface albedo

[2] It must be noted that the sigmoïd function of the optical depth only satisfies the boundary value problem of $C = 0$ for $\tau = 0$ asymptotically.

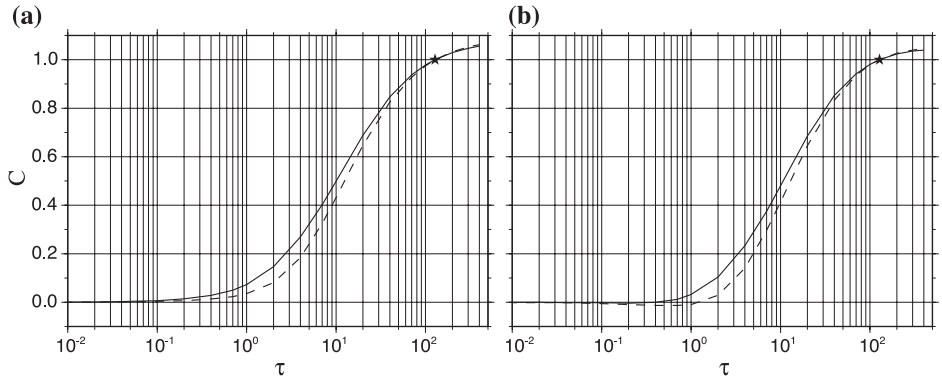

Fig. 1. Mean cloud amount C versus cloud optical depth τ empirical curves for (a) ocean surface albedo ($\alpha = 0.05$) under water cloud and (b) land surface albedo ($\alpha = 0.4$) under ice cloud using the STREAMER RT code. The plain curve is associated to $\theta_0 = 10°$, $\theta = 60°$ and $\varphi = 0°$, while the dashed curve is for $\theta_0 = 45°$, $\theta = 15°$ and $\varphi = 180°$. The stars (★) represent the optically opaque cloudy conditions ($\tau = 128$) above the surfaces.

and cloud phase instead of twice (C and τ) the number of optical depth values (here and according to Table 1 this number is equal to 15). Moreover, the relation between C and τ can be seen as an implicit non-linear interpolation on the discrete τ and C values computed by the RT code. We can directly see the benefit of the simple formulation of our fitting function. Indeed, Eq. (2) can explicitly be solved for the cloud optical depth, leading us to

$$\tau = \frac{\tau_0 C^{\chi}}{(A - BC)^{\chi}}, \tag{3}$$

with the mathematical constraints of Eq. (2) ($\chi \neq 0$, $\tau_0 > 0$) and setting the retrieved τ value to 128 when $C \geq A/B$ (saturation condition).

Thus, our cloud optical depth retrieval algorithm immediately follows. For each imager pixel:

(1) We estimate its cloud phase from the 0.67 and 1.6 μm SEVIRI channels according to Jolivet and Feijt (2003),
(2) According to the clear-sky reflectance estimated as in Ipe et al. (2003) and the pixel's viewing geometry (θ_0, θ, φ), we search the associated broadband surface albedo $\tilde{\alpha}$ corresponding to the closest match between the RT clear-sky imager simulated reflectances $\rho(\theta_0, \theta, \varphi, \tilde{\alpha}, \tau = 0)$ and this estimated clear-sky value,
(3) We compute its associated mean cloud amount value, C using Eq. (1), where the optically opaque cloudy reflectance, $\rho(\theta_0, \theta, \varphi, \tilde{\alpha}$, phase, $\tau = 128)$ is from RT simulation according to the pixel's viewing geometry (θ_0, θ, φ) and broadband surface albedo $\tilde{\alpha}$ found in the previous step,
(4) We compute its associated cloud optical depth value using Eq. (3) and the four fitted parameters for the associated pixel's viewing geometry, broadband surface albedo $\tilde{\alpha}$ and cloud phase.

2.2. Cloud fraction

The cloud fraction or cloud cover f is always defined on some footprint (i.e. a set of pixels) and simply consists in the computation of the relative fraction of cloudy pixels within that footprint. Major cloud detection algorithms found in literature are based on threshold decision tests related to multispectral radiance measurements for each pixel (see for example Rossow and Garder, 1993 or Saunders, 1986). Moreover, such a cloudy/non-cloudy test is generally applied before any cloud properties retrieval scheme. For GERB, we adopt a different approach. The cloud screening scheme requires that the cloud optical depth retrieval scheme described above has been priorly applied. The pixel is then flagged as cloudy if $\tau > \tau_{thres}$, where τ_{thres} is a constant threshold value (see Section 4). As mentioned in Section 1, we are planning to deliver TOA broadband fluxes at an increased spatial resolution of 10 km at nadir compared to the lower native GERB resolution. Therefore, due to the spatial resolution of SEVIRI, the footprint size on which cloud fraction needs to be estimated is 3×3 SEVIRI pixels.

3. Data description

To uncover possible discrepancies between GERB and CERES cloud properties retrieval methods and suggest some correction schemes, we compare the cloud optical depth and cloud fraction predicted by both instruments' algorithms. However, during the commissioning period of MSG, SEVIRI measurements will not be available on an operational basis (i.e. real-time dissemination through receiving station). Therefore, the RGP will be tested with Meteosat-7 as the imager. Nevertheless, application of the GERB SI scheme on MS-7 visible narrowband data and comparisons with CERES will prove useful for the validation of our adopted methodology for cloud optical properties retrievals.

Due to the lack of a 1.6 μm channel on MS-7, the cloud phase algorithm cannot be directly transposed to this imager. Note that we could use a cloud phase discrimination method based on brightness temperatures estimated from the MS-7 11.5 μm thermal channel, but this would require a further validation step. Therefore, as such a scheme would certainly not be used for SEVIRI due to its uncertainties compared to its CERES counterpart, we will rely on the determination of the cloud phase given in the CERES products.

Considering the relatively broad visible channel of MS-7 extending up to 1.0 μm compared to the narrower 0.6 and 0.8 μm SEVIRI bands (see Fig. 2), we are expecting an increase in the amount of radiation absorbed with respect to the cloud particle size as illustrated in Jolivet and Feijt (2003) for the SEVIRI 1.6 μm channel. Therefore, it will certainly introduce some scattering in GERB retrievals due to our adopted methodology (RT computed radiances were made with a fixed effective particle size for each cloud phase). *Rayleigh* and aerosol scattering as well as surface albedo tend to increase clear-sky measurements made in this wider visible channel, thus leading to a decrease in sensitivity of the radiance towards cloud optical depth. However, for SEVIRI, our cloud identification scheme will be applied for ocean on 0.8 μm data for minimal sensitivity regarding the

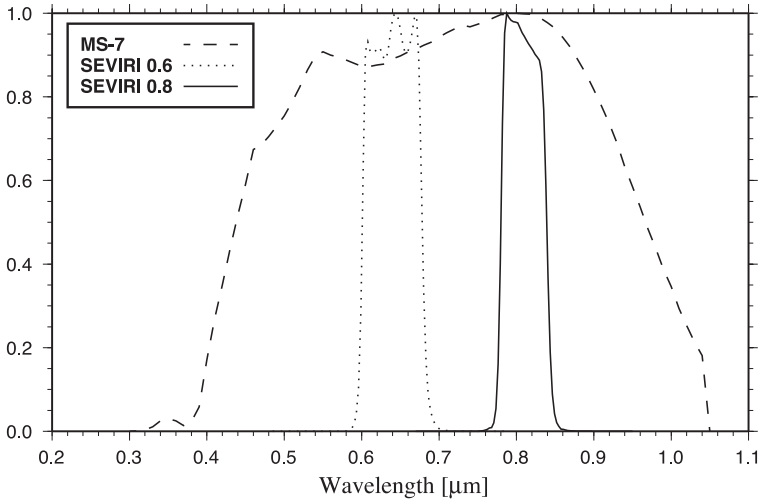

Fig. 2. Meteosat-7 and SEVIRI visible narrowband spectral responses.

surface albedo, *Rayleigh* and aerosol scattering. For other surface types the 0.6 μm channel will be considered in order to minimize the contribution of the surface albedo in the signal.

Since CERES ADMs were built from the Tropical Rainfall Measuring Mission (TRMM) satellite measurements using the Visible InfraRed Scanner (VIRS) imager, we shall use the CERES Single Satellite Footprint (SSF) TOA/Surface Fluxes and Clouds products which are generated on an hourly basis. Indeed, this dataset was used for the ADMs production (Loeb et al., 2003). As the operational phase of MS-7 started in the beginning of June 1998, we used for our comparisons 5000 × 5000 MS-7 visible images (about 2.25 km at nadir) from the Meteorological Archive and Retrieval Facility (MARF) of EUMETSAT and CERES SSF data from June 1998 to the end of August 1999 (15 months). Due to an anomaly of the CERES radiometer at the end of August 1998, we priorly use SSF VIRS Edition 2A and VIRS-only Edition 2 a posteriori. However, this change of data has no influence on the cloud parameters retrieval since CERES algorithms are only relying on VIRS data which were still available after the discovery of the anomaly.

To ensure measurements over similar scenes by both satellites (cloud motion due to the wind), our comparisons only take into account CERES footprints and related MS-7 pixels measured within 5 min around CERES acquisition time. Moreover, in order to limit sensitivity of our retrievals to parallax effects and 3-dimensional cloud structures, we only select for our comparisons CERES footprints that have a maximum tilt angle of 5° between the vectors joining the ground surface to each satellite. By this way, we select close scene viewing geometries for both instruments. In addition, only CERES footprints which are completely characterized by the VIRS imager are considered (i.e. the imager coverage within each footprint is 100%). Then, for each selected CERES footprint, we compute the mean cloud optical depth and cloud fraction over an equivalent GERB footprint made up of the 5 × 5 MS-7 nearest pixels to the geographical center of the

CERES footprint. As previously mentioned, we rely on the CERES SI scheme for the cloud phase determination. However, as this feature is an average over the CERES footprint, we do not have access to its associated values at the imager pixel-scale. Therefore, we only consider footprints with pure water or ice clouds according to CERES. Finally, to be able to easily bin our results according to the surface type, we further refine our selection of CERES footprints to those characterized by homogeneous geotypes.

4. Comparisons of GERB and CERES cloud properties

As we have seen, the mean cloud optical depth and cloud fraction within each GERB footprint are both linked to the chosen threshold τ_{thres} used to distinguish between cloudy and non-cloudy imager pixels. It can be shown that variations of this parameter mainly affect low cloud optical depth retrievals while high τ values are nearly insensitive. Requiring that the value of this parameter is taken in order that, in average on the studied footprints, the cloud fractions computed by the CERES algorithm are nearly identical to those computed by our scheme guarantees that we detect in average the same amount of cloud coverage than CERES. Practically, we find that a value of $\tau_{thres} = 0.85$ matches our requirement regarding the mean cloud coverages for both instruments as it is shown in Fig. 3.

To be consistent with the CERES methodology and thus to be able to perform meaningful comparisons between CERES and GERB results, we should convolve the cloud optical properties from imager pixels within GERB footprints with the dynamical CERES point spread function (PSF) as it is done for the retrievals in CERES SSF data. However, due to the increase of programming complexity and computing time, we have

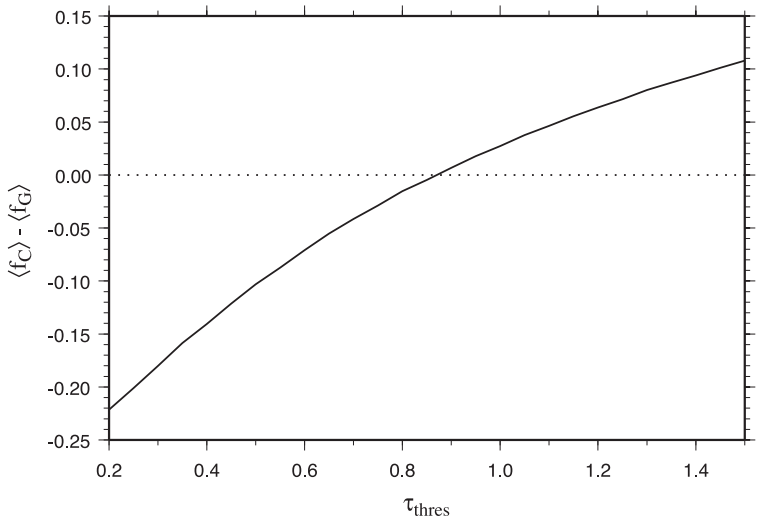

Fig. 3. Difference between CERES and GERB mean fractional cloud cover versus τ_{thres}.

discarded these calculations. Nevertheless, it is expected that such an approximation is not introducing any systematical bias but only random errors.

4.1. Cloud optical depth

The average of cloud optical depth over the equivalent GERB footprint is performed by a natural logarithmic mean on cloudy pixels based on the threshold value τ_{thres}. The range of possible τ values is spreading across three decades, while ln τ is varying almost linearly. This methodology is also adopted in the CERES products. Moreover, to avoid any bias when comparing the cloud optical depths, we only take into account CERES footprints having a single cloud layer as flagged so by the CERES SI scheme and corresponding to totally overcast cloudy conditions within the CERES field-of-view (FOV).

Despite the huge set of data used, the selection criteria on the cloud layers, ground surfaces and instruments tilt angle applied for the cloud optical depth comparisons only produce a limited set of pairs of values τ_C and τ_G (see Table 2). This table lists the different populations according to their CERES surface type and cloud phase. The most noticeable fact is the limited measurements above desert surface compared to others. This is not so surprising since deserts are known to be dry areas, and thus with very low occurrences of clouds (which are the media for precipitation).

The comparisons between CERES and GERB cloud optical depth retrieval schemes are illustrated in Fig. 4. For each panel in this figure we chose to plot the natural logarithm of the retrieved τ values in order to display small and large τ with the same order of magnitude. This comparison accounts for collocated footprints associated to nearly time-simultaneous measurements and similar viewing geometries of both imagers. The most noticeable result from these scatterplots is the good agreement between the two retrieval schemes. The correlation coefficients are 0.894 for ocean with water clouds, 0.946 for ocean with ice clouds, 0.874 for vegetation with water clouds, 0.924 for vegetation with ice clouds and 0.886 for desert with ice clouds, but only 0.576 for water clouds above desert (see Fig. 4(e)). This can be explained by the fact that the associated dataset population is badly distributed and thus not statistically significant. The same conclusion reaches to the validity of the linear least square fit for that case. Regarding ice clouds cases, there is a larger scatter of the plots compared to the water phase cases as illustrated by the χ^2 values of the linear fits (as for example 0.094 versus 0.106 for ocean). A possible explanation of this effect could be related to the broader distribution of the ice particle effective radii r_e for ice clouds than for water clouds (see Fig. 5). One can argue that using

Table 2
Population of GERB and CERES cloud optical depth pairs according to their associated CERES surface type and cloud phase

	Water	Ice	Total
Ocean	12646	3645	16291
Vegetation	768	1770	2538
Desert	116	187	303
Total	13530	5602	19132

Only pure footprints according to the CERES mean cloud phase with homogene surface are considered.

visible radiance measurements, the impact of the ice particle size distribution on the cloud optical depth retrievals should normally be less than the one we show. In fact, we strongly question our assumption regarding the insensitivity of visible radiance to cloud particle size when using MS-7 due to its broad visible channel (see Fig. 2) which extends at least partly in the near-infrared region. Moreover, this could also explain the deviation of the linear least square fits relative to the ideal case line.

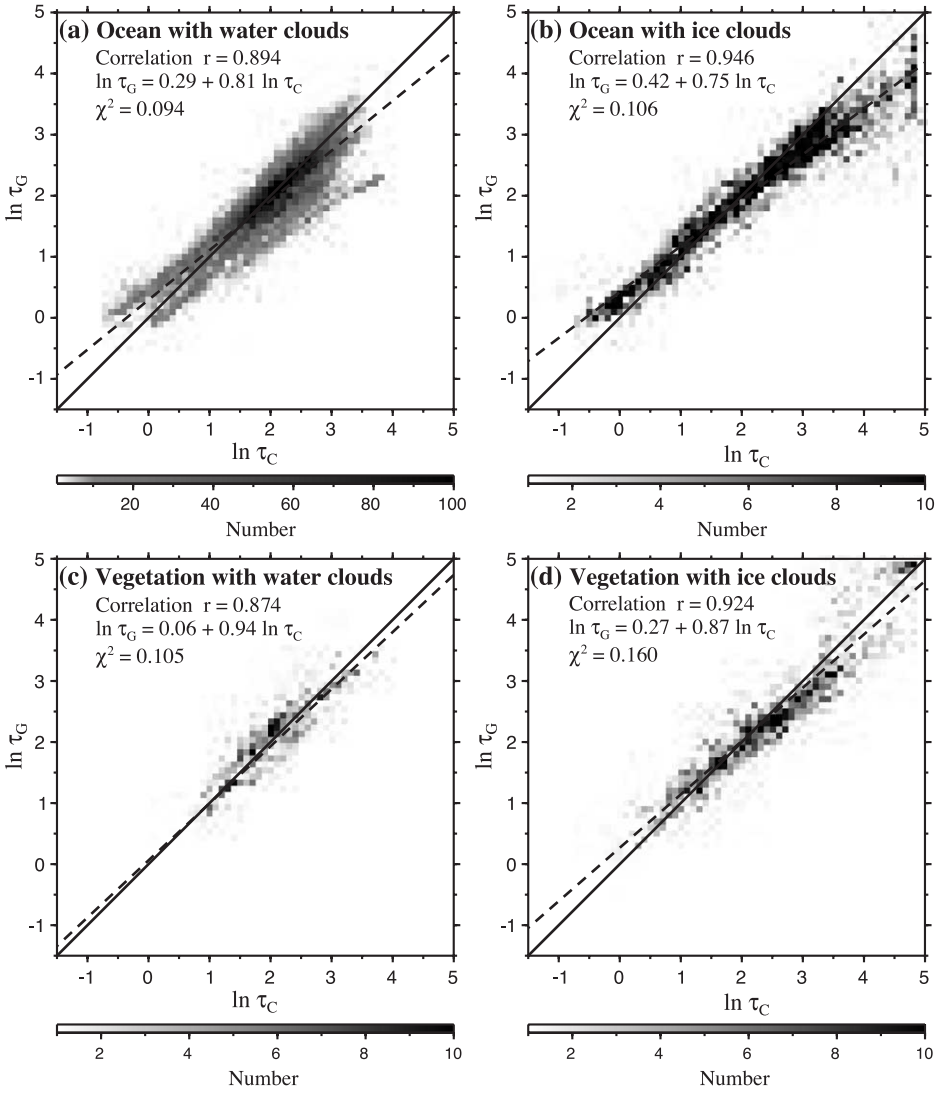

Fig. 4. Natural logarithm of cloud optical depth retrieval pairs ($\ln \tau_C$, $\ln \tau_G$) from collocated CERES and GERB footprints according to the different considered surfaces and cloud phases. Plain line represents the ideal case $\tau_G = \tau_C$ while dashed line illustrates a robust linear least square fit with respect to outliers (Huber, 1964).

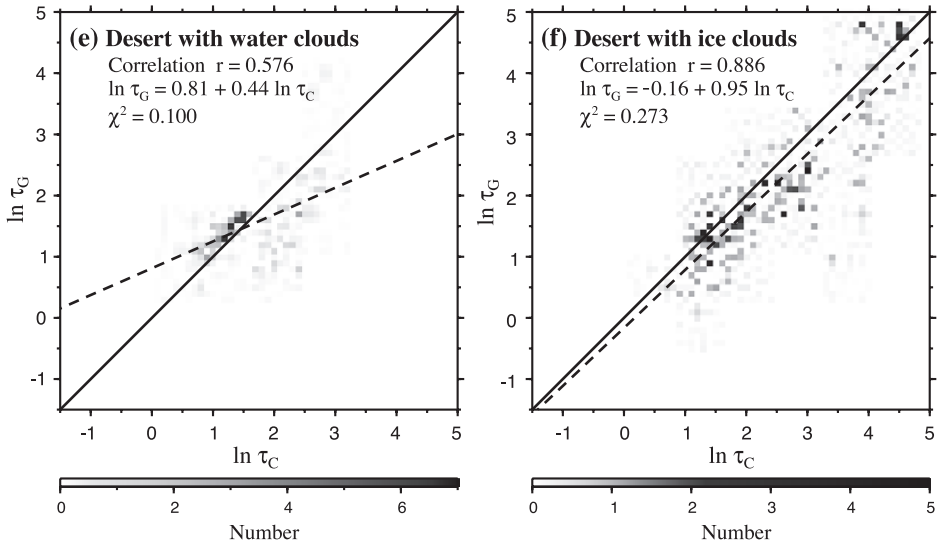

Fig. 4 (*continued*).

As mentioned previously, the implementation of a scene identification within the RGP is used to select the proper CERES ADMs for the broadband radiance-to-flux conversions. These ADMs are defined over some ranges of cloud optical depth (Loeb et al., 2003). Hence, the crucial constraint of our retrieval scheme is to select in average the same ADM as the CERES software would (i.e. having a confusion matrix between the two classifications as close as possible diagonal). Such a confusion matrix is shown in Table 3 for ocean and ice cloud footprints.[3] As we could expect from the correlation between ln τ_C and ln τ_G (Fig. 4(b)) retrievals are located around its diagonal while spreading away with increasing cloud optical depth ranges. However, due to the deviation from the ideal line of the least square fit, the sub-matrices associated to extreme cloud optical depth values are not symmetric with all their diagonal elements above any other corresponding row- or column-elements (diagonal dominance). If such requirements were fulfilled, GERB cloud optical depth retrieval scheme would lead on average to a similar selection of ADMs as CERES would (diagonal dominance) and misselections of upper and lower ADMs would occur at similar frequencies (symmetric matrix). However, these previous results suggest the need to develop some correction strategy in order to adjust GERB retrievals to CERES ones. Such a scheme will be developed later in this paper.

4.2. Cloud fraction

To avoid any bias when comparing the cloud fraction, we only take into account CERES footprints having a single cloud layer as flagged so by the CERES SI scheme (there is no restriction on cloud coverage). The population of footprints matching this

[3] Confusion matrices for Fig. 4(a), (c)–(f) will not be shown to spare space.

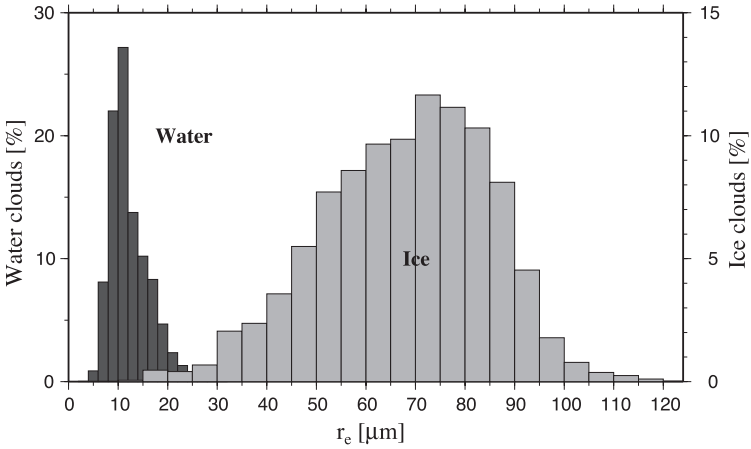

Fig. 5. Normalized histograms of the CERES cloud effective particle size for the selected footprints used in the cloud optical depth comparisons.

criterion is given in Table 4 according to the surface type and cloud phase. Due to the fact that we did not perform CERES PSF convolution of our retrievals on SEVIRI pixels within GERB footprints, GERB cloud fraction values are only part of the discrete set made of the multiples of $1/(5 \times 5) = 0.04$, while they are continuously varying for CERES. To be able to perform meaningful comparisons, we have plotted in Fig. 6 both cloud fractional

Table 3
Confusion matrix (in percent) resulting from the comparisons of CERES ADMs cloud optical depth classifications according to CERES (columns) and GERB (rows) retrievals for ocean and ice cloud footprints

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	**0.00**														
2		**0.91**	0.08												
3		4.77	**10.32**	1.37	0.03										
4		0.03	2.85	**9.33**	1.32	0.27	0.08	0.03							
5				2.39	**4.14**	1.67	0.69	0.19	0.03		0.05	0.03			
6				0.44	1.95	**2.66**	1.43	0.44	0.27	0.22	0.22	0.11	0.05	0.11	0.49
7					0.33	0.88	**1.59**	1.62	1.15	0.52	0.27	0.25	0.08	0.14	0.41
8						0.14	0.96	**1.37**	1.23	0.82	1.18	0.25	0.55	0.25	0.60
9						0.05	0.16	0.60	**0.85**	1.07	1.26	0.93	0.66	0.19	0.49
10							0.05	0.14	0.44	**0.55**	0.99	0.71	0.91	0.36	0.66
11							0.03	0.03	0.22	0.77	**1.23**	1.45	1.59	0.91	1.48
12									0.05	0.14	0.33	**0.93**	1.04	0.71	2.11
13										0.03	0.08	0.36	**0.74**	0.69	2.83
14												0.05	0.36	**0.16**	2.50
15										0.03	0.05	0.03	0.25	0.19	**3.98**

The bold numbers on the diagonal are for cases where the ADMs selected using CERES and GERB cloud optical depth retrievals are identical. ADM class numbers are for the following ranges of cloud optical depth (Loeb et al., 2003): 1=(0–0.01), 2=(0.01–1), 3=(1–2.5), 4=(2.5–5), 5=(5–7.5), 6=(7.5–10), 7=(10–12.5), 8=(12.5–15), 9=(15–17.5), 10=(17.5–20), 11=(20–25), 12=(25–30), 13=(30–40), 14=(40–50) and 15=(>50).

Table 4
Population of GERB and CERES cloud optical depth pairs according to their associated CERES surface type and cloud phase

	Water	Ice	Total
Ocean	40599	4439	45038
Vegetation	4055	2008	6063
Desert	1507	430	1937
Total	46161	6877	53038

Only pure footprints according to the CERES mean cloud phase with homogene surface are considered.

cover retrievals for ocean footprints. These results have been binned in squared regions according to a bin width of 0.04. It can be noted that there is a good agreement for clear-sky and overcast footprints between both schemes and that these cases are a major contribution to the total population of selected footprints (notice the logarithmic scale of the colorbars in the graphs). Moreover, according to the gray band joining these extreme cloudy conditions in Fig. 6(a), there is also a correlation for intermediate cloud fraction values.

However, results are also more scatter than for the cloud optical depth. It can be explained by two joint contributions. A horizontal scatter is due to the fact that CERES retrievals are convoluted with its dynamical PSF which is stretching around 72 VIRS pixels (Smith, 1994) and thus taking into account a larger footprint than the expected 5×5 VIRS pixels when estimating cloud fraction. This is clearly illustrated in Fig. 6 where $f_G = 0$ and $f_G = 1$ while $f_C \in (0-1)$. While τ_{thres} has only a limited impact on the low cloud optical depth retrievals, it is a highly sensitive parameter for cloud fraction estimation (which is totally relying on the cloudy/non-cloudy test). On the opposite, CERES methodology based on multispectral radiance threshold tests is not directly depending on the retrieved cloud optical depths. This can therefore result in cloudy footprints for which the mean cloud optical depth is below than the GERB chosen threshold. This contributes to the horizontal scatter and also to its vertical component as shown in the figure where $f_C = 1$ while $f_G \in (0-1)$. But we can question the validity of such low CERES cloud optical depth retrievals with regards, for example, to the sensitivity in the estimation of the surface albedo.

As the CERES ADMs are also stratified according to some ranges of the cloud fraction (Loeb et al., 2003), the confusion matrix for ocean and water cloud footprints is given in Table 5.[4] From this table, it can be noted that while about 34% of the footprints are identified as totally overcast by CERES, it reaches up to 45% for GERB. Moreover, GERB is also overestimating the number of footprints in the clear-sky class (15%) compared to CERES (0.42%). In fact, it can be noticed that the majority of these GERB retrievals are classified in CERES class 2. This is certainly due to the range used to defined classes 1, 2, 12 and 13 on extremes cloud fraction values and to the convolution of CERES retrievals which tend to spread their values compared to GERB results as well as probable CERES cloud optical depths which are below our chosen threshold. As previously mentioned, a possible solution to correctly compare CERES and GERB cloud fractional

[4] Confusion matrices for other surfaces and phase will not be shown to spare space.

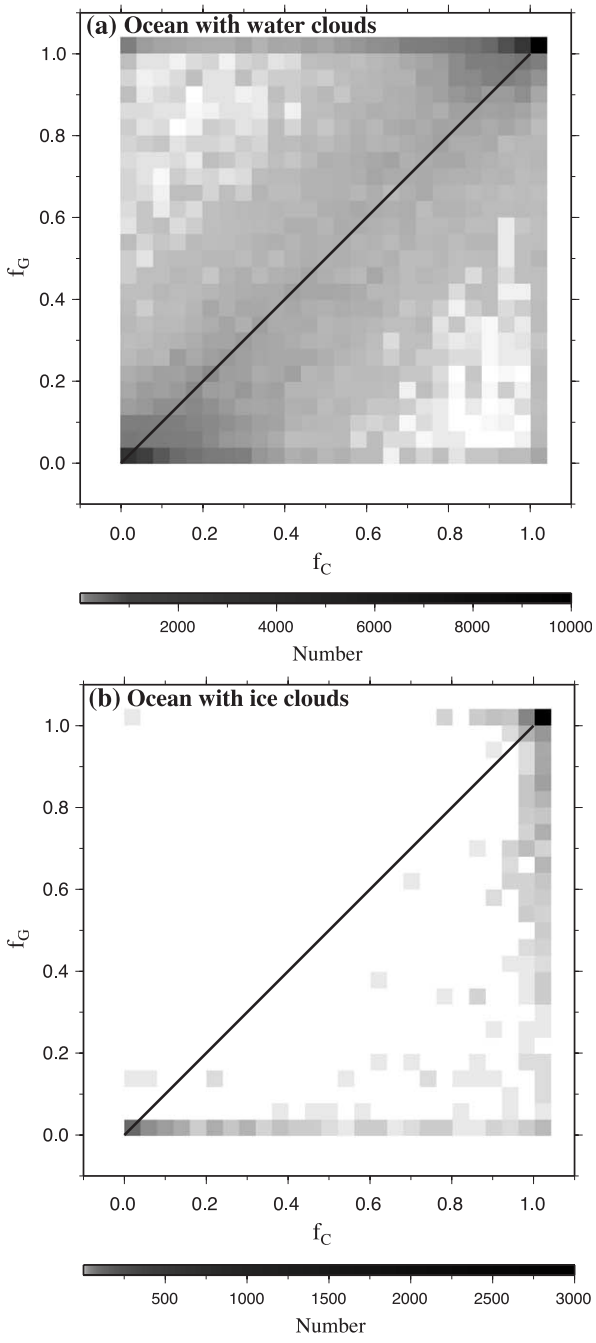

Fig. 6. Cloud fraction retrieval pairs (f_C, f_G) from collocated CERES and GERB ocean footprints according to both cloud phases. Plain line represents the ideal case $f_G = f_C$ where both retrieval schemes provide identical values.

Table 5
Confusion matrix (in percent) resulting from the comparisons of CERES ADMs cloud fraction classifications according to CERES (columns) and GERB (rows) retrievals for ocean and water cloud footprints

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	Sum
1	**0.40**	10.51	2.41	0.97	0.42	0.21	0.14	0.07	0.05	0.04	0.04	0.03	0.11	15.41
2	0.01	**1.86**	1.41	0.93	0.62	0.24	0.16	0.09	0.05	0.02	0.01	0.01	0.09	5.50
3		0.68	**0.96**	0.76	0.59	0.34	0.20	0.13	0.08	0.04	0.02	0.03	0.07	3.89
4	0.00	0.47	0.82	**1.07**	0.84	0.67	0.40	0.21	0.15	0.08	0.02	0.03	0.10	4.88
5		0.18	0.23	0.47	**0.51**	0.44	0.35	0.23	0.13	0.08	0.03	0.04	0.07	2.78
6		0.18	0.31	0.41	0.52	**0.69**	0.60	0.47	0.29	0.19	0.07	0.06	0.11	3.90
7		0.11	0.12	0.18	0.29	0.38	**0.38**	0.33	0.28	0.22	0.06	0.05	0.06	2.46
8		0.11	0.14	0.16	0.33	0.43	0.57	**0.65**	0.56	0.42	0.17	0.15	0.11	3.79
9	0.00	0.08	0.07	0.10	0.12	0.21	0.35	0.43	**0.51**	0.45	0.14	0.12	0.11	2.71
10		0.12	0.06	0.09	0.14	0.19	0.36	0.56	0.77	**1.04**	0.47	0.49	0.23	4.52
11		0.03	0.04	0.04	0.04	0.05	0.08	0.16	0.29	0.52	**0.32**	0.30	0.18	2.06
12		0.05	0.03	0.03	0.04	0.06	0.09	0.16	0.31	0.60	0.37	**0.69**	0.26	2.68
13	0.01	0.46	0.26	0.24	0.23	0.29	0.36	0.53	0.81	1.61	1.85	6.08	**32.68**	45.42
Sum	0.42	14.86	6.87	5.46	4.69	4.20	4.05	4.03	4.27	5.33	3.57	8.08	34.17	40 599

Bold numbers on the diagonal are for cases where the ADMs selected using CERES and GERB cloud fraction retrievals are identical. ADM class numbers are for the following ranges of cloud fraction (Loeb et al., 2003): 1=(0–0.001), 2=(0.001–0.1), 3=(0.1–0.2), 4=(0.2–0.3), 5=(0.3–0.4), 6=(0.4–0.5), 7=(0.5–0.6), 8=(0.6–0.7), 9=(0.7–0.8), 10=(0.8–0.9), 11=(0.9–0.95), 12=(0.95–0.99) and 13=(0.99–1).

covers would be to perform CERES PSF convolution on MS-7 pixels and only select footprints whose mean cloud optical depth is above τ_{thres}. However, such processes are deferred to a future study on SEVIRI data. Nevertheless, the choice of τ_{thres} is a basic validation of our cloud fraction retrievals as it leads to the same average of cloud fractional cover than CERES.

5. Homogenization of GERB cloud optical depths

5.1. Method

In this section, we propose a generic homogenization scheme to correct retrievals of a given feature according to some method towards associated reference values. More specifically, the reference values will be taken as the CERES cloud optical depth retrievals while the results that have to be corrected are the outputs of the GERB cloud optical depth algorithm.

As previously mentioned, cloud optical depth values are not convenient to handle due to their large range of variation (about three decades). Any attempt to fit them directly will produce poor results due to the influence of large values compared to small ones. One would expect that retrieval errors increase with increasing optical depth values and thus, any fit should decrease the weights of these large values accordingly. However, if we consider the logarithm of cloud optical depths, the variation range is implicitly reduced while having an almost linear dependence and therefore these values are more suited to perform meaningful statistical fits.

To perform statistical fit on ln τ values we used by convenience a mathematical model f (needed to express the correlation between two variables) of the following form: $f(\ln\tau) = \mathcal{P}_n(\ln\tau)$, where $\mathcal{P}_n$ is the generic form of a degree n polynomial with coefficients a_i for $i=0, \ldots, n$. The robust polynomial least square fits with respect to outliers (Huber, 1964) which lead to a significant decrease of the residuals for the selected footprints are given in

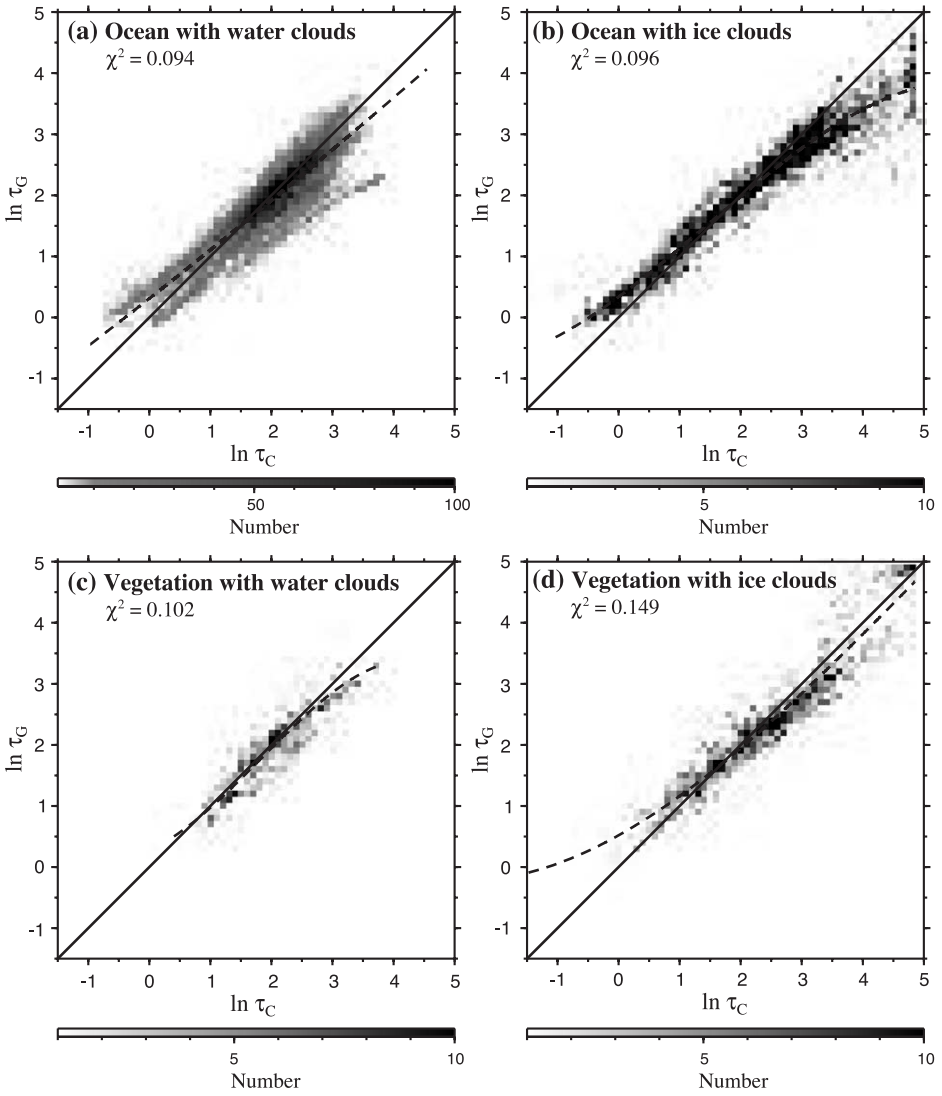

Fig. 7. Natural logarithm of cloud optical depth retrieval pairs (ln τ_C, ln τ_G) from collocated CERES and GERB footprints according to three different surfaces and two cloud phase types. Plain line represents the ideal case $\tau_G = \tau_C$ where both retrieval schemes provide identical values. Dashed line illustrates the most significative robust polynomial least square fits in terms of confidence on the decrease of residuals (Huber, 1964).

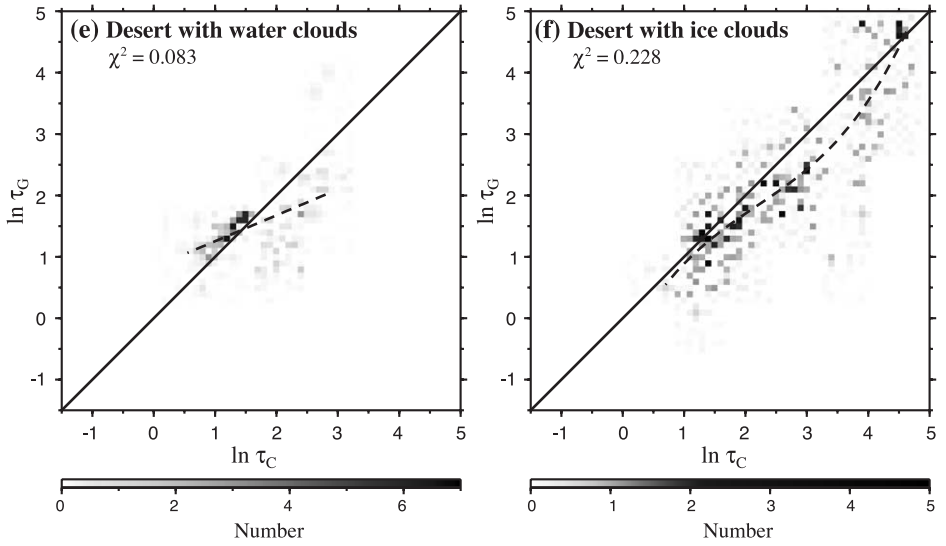

Fig. 7 (*continued*).

Fig. 7. Having set the number of degrees of freedom n^* and performed the fit on the dataset, we end up with the expression

$$\ln\tau_G = \mathcal{P}_{n^*}(\ln\tau_C) \tag{4}$$

where $\mathcal{P}_{n^*}$ is the fitted polynomial of degree n^* with coefficients a_i^* for $i=0,\ldots,n^*$. The next step is to homogenize values for GERB cloud optical depth, τ_G, in such a way that in average $\ln\tau_C = \ln\tau_G$. By substituting the latter formula in Eq. (4), the homogenized cloud optical depths are implicitly given by

$$\mathcal{P}_{n^*}(\ln\tilde{\tau}_G) = \ln\tau_G. \tag{5}$$

As we see, the difficulty lies in solving this equation for $\tilde{\tau}_G$, either explicitly for $n^*\leq 5$ or iteratively when $n^*>5$, for every value τ_G. However, robust iterative methods for finding roots of polynomials exist and are described in the literature (Press et al., 1992).

5.2. Results

Our correction scheme is illustrated for cloud optical depth pairs associated with ocean and ice cloud footprints. It turns out that the fitting model in terms of robustness and significant decrease of the residuals is the third degree polynomial $P_4(\ln\ \tau_C) = 0.336 + 0.737\ln\ \tau_C + 0.076\ (\ln\ \tau_C)^2 - 0.017\ (\ln\ \tau_C)^3$. Results are given in Table 6 as a confusion matrix. As we could expect from Fig. 4(b), our retrieval scheme overestimates low cloud optical depth while underestimates high τ values by comparison to CERES. The benefit of our correction method is clearly shown by comparing uncorrected (Table 3) and homogenized GERB values (Table 6) where the number of

Table 6
Confusion matrix (in percent) resulting from the comparisons of CERES ADMs cloud optical depth classifications according to CERES (columns) and GERB (rows) retrievals for ocean and ice cloud footprints for homogenized GERB results

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	**0.00**														
2		**4.61**	1.73	0.03											
3		1.07	**9.96**	2.55	0.05										
4		0.03	1.56	**8.64**	1.54	0.33	0.08	0.03							
5				1.81	**3.68**	1.48	0.63	0.19	0.03		0.05	0.03			
6				0.49	1.78	**2.33**	0.91	0.22	0.19	0.08	0.11	0.05	0.05	0.11	0.25
7					0.63	0.93	**1.45**	0.85	0.47	0.30	0.11	0.19		0.05	0.36
8					0.08	0.47	0.96	**1.10**	0.85	0.47	0.36	0.11	0.14	0.11	0.41
9						0.05	0.69	0.99	**0.77**	0.55	0.69	0.16	0.30	0.14	0.30
10						0.05	0.14	0.47	0.91	**0.47**	0.91	0.41	0.33	0.16	0.44
11						0.03	0.11	0.55	0.55	1.07	**1.43**	0.91	0.96	0.16	0.49
12							0.03		0.33	0.49	0.69	**0.74**	0.85	0.41	0.69
13								0.03	0.14	0.55	0.93	1.21	**1.84**	0.85	1.23
14									0.03	0.14	0.25	0.69	0.69	**0.66**	1.54
15										0.03	0.16	0.60	1.07	1.04	**9.85**

The bold numbers on the diagonal are associated to the case where the ADMs selected using CERES and GERB cloud optical depth retrievals are identical. ADM class numbers are for the following ranges of cloud optical depth (Loeb et al., 2003): 1=(0–0.01), 2=(0.01–2.5), 3=(2.5–6), 4=(6–10), 5=(10–18), 6=(18–40), 7=(>40).

diagonal-dominant elements decreased from 8 to 4 while the off-diagonal elements are balanced, leading to a more symmetric matrix. Finally, it can be noted that adjusted GERB values are now all classified in an adjacent CERES class for those that do not meet the diagonal dominance criterion.

6. Conclusions and future work

In this paper, we have presented cloud optical depth and cloud fraction retrieval algorithms. These schemes will be implemented as part of the SEVIRI scene identification for the RMIB GERB Ground Segment. They are based on a direct inversion method that relies on simulated radiances by a radiative transfer model over specific scenes, composed by a restricted number of ground surfaces and cloudy conditions. The method uses an innovative and efficient parameterization of the inversion look-up tables. Preliminary validations of the cloud optical depth and cloud fraction retrieval algorithms were performed by comparing CERES footprints and Meteosat-7 simulated GERB footprints. These comparisons have shown good correlation between GERB and CERES methods. However, discrepancies occur but there are expected to decrease with the use of the SEVIRI imager thanks to its narrower visible channels compared to Meteosat-7. But it will have to be confirmed by a future validation study on this imager. Moreover, a correction scheme was also suggested in order to cope with the possible remaining cloud optical depth discrepancies between CERES and SEVIRI retrievals. Its efficiency was demonstrated on CERES and MS-7 cloud optical depth comparisons. Basic validation of the cloud fraction was performed by choosing the optimal threshold value on cloud optical

depth which is the border between cloudy/non-cloudy imager pixels. This value ensures that, in average, our algorithm produces the same cloud coverage as the one computed by the CERES software. Finally, some recommendations were issued for future comparisons such as convolute with CERES point spread function our cloud properties retrievals.

These cloud optical depth and cloud fraction estimations will be part of the GERB products generated at RMIB. Complete description of their content as well as the algorithms used are available at http://gerb.oma.be. Products will be freely available to the scientific community at ftp://gerb.oma.be.

Acknowledgements

The CERES SSF data were obtained from the NASA Langley Research Center Atmospheric Sciences Data Center. The Meteosat-7 data were obtained from the Meteosat Archive and Retrieval Facility (MARF). All graphics and plots were made using The Generic Mapping Tools (GMT) softwares collection of P. Wessel (University of Hawai'i) and W.H.F. Smith (NOAA/NESDIS/NODC).

References

Arking, A., Childs, J.D., 1985. Retrieval of cloud cover parameters from multispectral satellite images. J. Clim. Appl. Meteorol. 24, 322–333.

Cano, D., Monget, J.M., Albuisson, M., Guillard, H., Regas, N., Wald, L., 1986. A method for the determination of the global solar radiation from meteorological satellite data. Sol. Energy 37, 31–39.

Dewitte, S., Clerbaux, N., Gonzalez, L., Hermans, A., Ipe, A., Joukoff, A., Sadowski, G., 2000. Generation of GERB unfiltered radiances and fluxes. Proc. 2000 EUMETSAT Meteorological Satellite Data Users' Conf., EUMETSAT, Bologna, pp. 619–625.

Gonzalez, L., Hermans, A., Dewitte, S., Ipe, A., Sadowski, G., Clerbaux, N., 2000. Resolution enhancement of GERB data. Proc. 2000 EUMETSAT Meteorological Satellite Data Users' Conf., EUMETSAT, Bologna, 619–625.

Harries, J.E., Crommelynck, D., 1999. The Geostationary Earth Radiation Budget experiment on MSG-1 and its potential applications. Adv. Space Res. 24, 915–919.

Huber, P.J., 1964. Robust estimation of a location parameter. Ann. Math. Stat. 35, 73–101.

Ipe, A., Bertrand, C., Clerbaux, N., Dewitte, S., Gonzalez, L., 2003. Pixel-scale composite top-of-the-atmosphere clear-sky reflectances for Meteosat-7 visible data. J. Geophys. Res. 108 (D19), 4612 (doi:10.1029/2002JD002771).

Jolivet, D., Feijt, A.J., 2003. Cloud thermodynamic phase and particle size estimation using the 0.67 and 1.6 μm channels from meteorological satellites. Atmos. Chem. Phys. Discuss. 3, 4461–4488.

Key, J., Schweiger, A.J., 1998. Tools for atmospheric radiative transfer: Streamer and FluxNet. Comput. Geosci. 24, 443–451.

King, M.D., 1987. Determination of the scaled optical thickness of clouds from reflected solar radiation measurements. J. Atmos. Sci. 44, 1734–1751.

Loeb, N.G., Smith, N.M., Kato, S., Miller, W.F., Gupta, S., Minnis, P., Wielicki, B.A., 2003. Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the tropical rainfall measuring mission satellite. Part I: Methodology. J. Appl. Meteorol. 42, 240–265.

Nakajima, T.Y., King, M.D., 1990. Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. J. Atmos. Sci. 47, 1878–1893.

Nakajima, T.Y., Nakajima, T., 1995. Wide-area determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEX regions. J. Atmos. Sci. 52, 4043–4059.

Press, W.H., Vetterling, W.T., Teukolsky, S.A., Flannery, B.P., 1992. Numerical Recipes in C. The Art of Scientific Computing, Second Edition. Cambridge University Press, Cambrige.

Rossow, W.B., Garder, L.C., 1993. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. J. Climate 6, 2341–2369.

Saunders, R.W., 1986. An automated scheme for the removal of cloud contamination from AVHRR radiances over western Europe. Int. J. Remote Sens. 7, 867–888.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., Ratier, A., 2002. An introduction to meteosat second generation (MSG). Bull. Am. Meteorol. Soc. 83, 977–992.

Smith, G.L., 1994. Effects of time response on the point spread function of a scanning radiometer. Appl. Opt. 33 (30), 7031–7037.

Suttles, J.T., Green, R.N., Minnis, P., Smith, G.L., Staylor, W.F., Wielicki, B.A., Walker, I.J., Young, D.F., Taylor, V.R., Stowe, L.L., 1988. Angular radiation models for the earth–atmosphere system. Shortwave Radiation, vol. I. NASA Ref. Publ RP-1184.

Wielicki, B., Barkstrom, B.R., Harrison, E.F., Lee, R.B., Smith, G.L., Cooper, J.E., 1996. Clouds and the Earth's Radiant Energy System (CERES): an earth observing system experiment. Bull. Am. Meteorol. Soc. 77, 853–868.

Wyser, K., O'Hirok, W., Gautier, C., Jones, C., 2002. Remote sensing of surface solar irradiance with corrections for 3-D cloud effects. Remote Sens. Environ. 80, 272–284.